

# Neuronal Correlates of View Representation Revealed by Face-View Aftereffect

Behrad Noudoost<sup>1</sup> and Hossein Esteky<sup>1,2</sup>

<sup>1</sup>School of Cognitive Sciences, Institute for Research in Fundamental Sciences, Tehran, Iran, and <sup>2</sup>Research Center for Brain and Cognitive Sciences, Shaheed Beheshti University of Medical Sciences, Tehran, Iran

Recognition of face-views is a crucial component of our social behavior. However, little is known about the neuronal basis of face-view recognition. Recent psychophysical studies have shown that adaptation to a particular face-view changes perception of other views in humans. To understand the neuronal basis of face-view representation and the mechanisms involved in face-view adaptation, we studied the responses of inferotemporal (IT) cortex neurons before and after face-view adaptation. We first used face-view adaptation to show that monkeys, as well as humans, exhibit face-view aftereffect, suggesting the presence of a view-sensitive representation in primates. Then, we recorded from IT neurons of monkeys and found that changes in responses of these neurons can account for changes in perception of face views attributable to adaptation, indicating IT as a neural correlate of face-view perception. More importantly, we provide evidence that IT neurons with wide-view tuning not only conveyed face-view information but also exhibited neural signatures of view aftereffect before neurons with narrower tuning. Our findings show that view-dependent representation of objects does not necessarily depend on narrow-tuned neurons. The finding that wide-view tuned neurons account for view-specific perceptual changes implies the importance of accumulating sensitivity and tolerance at the level of IT neurons, as the final stage of visual object recognition.

## Introduction

When looking at your friend's face, you can recognize her independent of whether she is facing toward or away from you. However, understanding the viewing direction of her face is also important in this situation, helping to manage social interactions. What are the neuronal bases of view-dependent and view-tolerant behaviors? Neurons in the inferotemporal (IT) cortex are involved in recognition of visual objects and faces (Gross et al., 1972; Perrett et al., 1992; Logothetis and Pauls, 1995; Logothetis and Sheinberg, 1996; Tanaka, 1996; Afraz et al., 2006). Whereas some IT neurons respond to a wide range of face views, others respond selectively to a more limited range. Some theories of object recognition suggest that the wide-range neurons are object-selective units that shape their view-invariant characteristics by pooling over activity of various neurons responding to narrower ranges of views. These models propose that view-sensitive neurons underlie our view-dependent behavior, whereas neurons responding to wider range of views are the substrates for our view-tolerant recognitions (Perrett and Oram, 1993; Logothetis and Pauls, 1995; Riesenhuber and Poggio, 1999). The goal of the current study is to characterize the neuro-

nal correlates of view-dependent object recognition by using a face-view adaptation paradigm. We hypothesized that, if IT neurons are the neuronal correlates of our view perception, they should change their response according to changes in perception attributable to face-view adaptation. It is known that the prolonged presentation of a particular face view (view-adaptation) to human subjects alters the perception of other views (view aftereffect) (Fang and He, 2005). For example, after adaptation to a particular view (e.g., 30° to the side), the perceived viewing angles of other views (e.g., 45°) are biased in the opposite direction toward larger angles (perceptual repulsive shift). In this study, we first showed the presence of face-view aftereffect in humans and monkeys and then used the adaptation-induced perceptual repulsive shift as a tool to search for neuronal correlates of face-view perception. We found that adaptation-induced changes in the responses of population of IT neurons can account for the observed view aftereffect.

As the third step in our search for neuronal correlates of view perception, we refined our method to assess the contribution of different functional classes of neurons in the view aftereffect. Comparing the signatures of view aftereffect between different groups of neurons, we found that the firing rate changes in neurons responding to a wide range of face views most closely resemble changes that occurred in face-view perception. Our results show that wide-tuned neurons play a role in view representation.

## Materials and Methods

### Human and monkey psychophysics

We trained two male monkeys to categorize images of 14 views of one artificial face identity into two different categories based on viewing direction of images (Fig. 1A). Views ranged from left profile (0°) to front view (90°) in equal steps, and subjects were asked to categorize them into

Received June 13, 2012; revised Feb. 8, 2013; accepted Feb. 18, 2013.

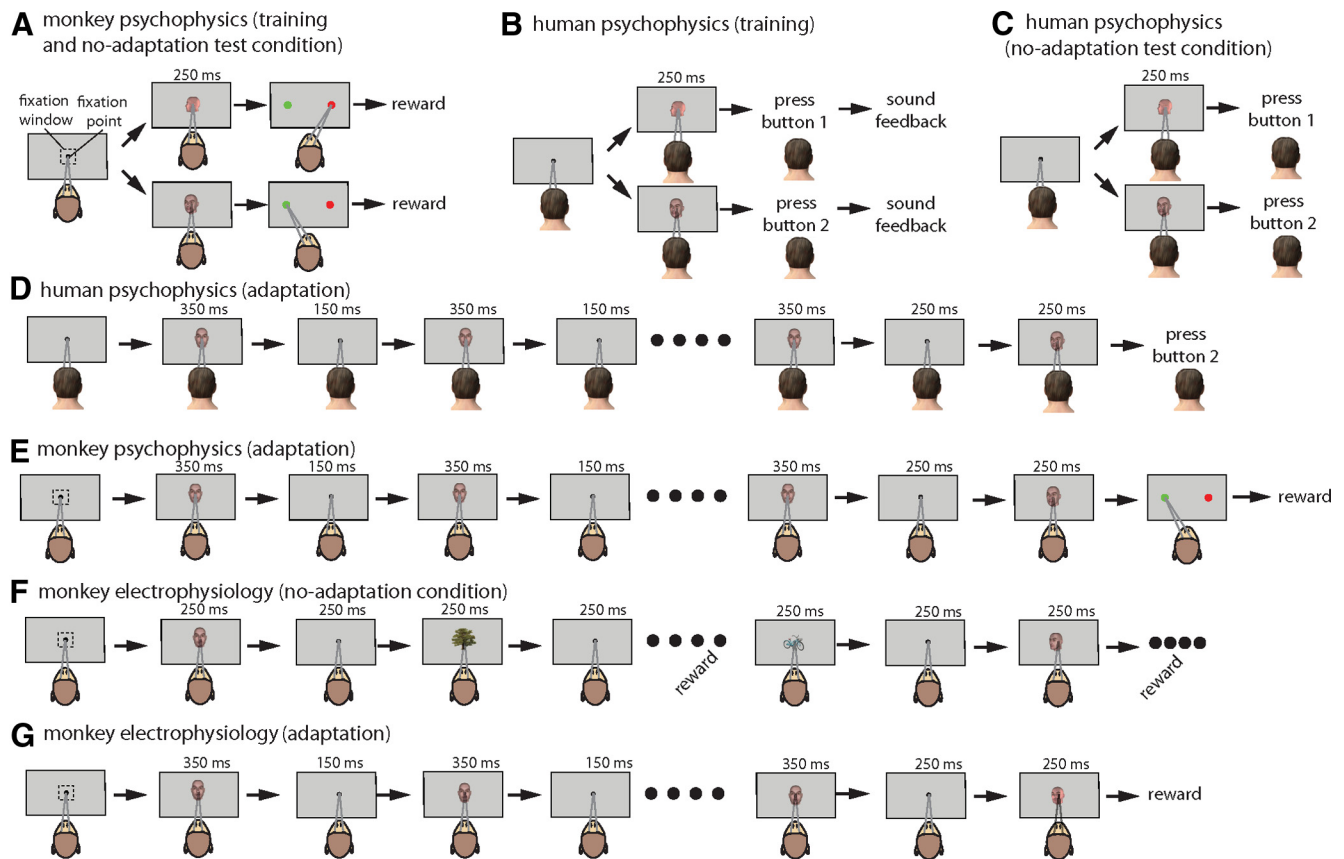
Author contributions: B.N. and H.E. designed research; B.N. performed research; B.N. analyzed data; B.N. and H.E. wrote the paper.

We thank David A. Leopold, Ali Asadollahi, Kelsey L. Clark, Ryan F. Squire, and Marc Zirnsak for valuable comments on this manuscript, Roozbeh Kiani for his comments on optimal decoding method, and Mehdi Adibi and Mehdi Sanayei for their help in human psychophysics part of the study. We also thank Neda Nategh for her help in designing the SVM classifier.

Correspondence should be addressed to Dr. Hossein Esteky, School of Cognitive Sciences, Institute for Studies in Fundamental Sciences, Niavaran, Tehran, Iran. E-mail: esteky@ipm.ir.

DOI:10.1523/JNEUROSCI.2838-12.2013

Copyright © 2013 the authors 0270-6474/13/335761-12\$15.00/0



**Figure 1.** Behavioral paradigms in psychophysics and electrophysiology parts of the study. **A**, In the monkey psychophysics part of the study, a trial would start when the monkey fixates within a fixation window of  $2^\circ \times 2^\circ$ . To receive a juice reward, the monkeys were required to choose the right (always red) target by moving their gaze after presentation of a left-profile view. Choosing the opposite target (always green on left) was rewarded in the case of front stimulus presentation. The monkeys were not rewarded for choosing the wrong target (e.g., saccading to red target after front-view stimulus). Throughout the training, the monkey was always rewarded for choosing the red target during presentation of stimuli  $<45^\circ$  (similar to profile view) or choosing the green one for  $>45^\circ$  stimuli. The test phase was completely similar to the end of the training phase, and the feedback (the juice) was still present. **B**, The training phase of human subjects was similar to the training phase for monkeys except that subjects were asked to press two different buttons for stimuli  $>45^\circ$  or  $<45^\circ$ . Feedback about the correct and wrong selections was made by presentation of two different sounds. Unlike monkeys, fixation of human subjects was not monitored. **C**, The test was similar to the training stage, except that subjects would not receive feedback about their correct or wrong selections. **D**, In adaptation conditions, everything was similar to the no-adaptation condition except that the presentation of the test stimulus was preceded by 5 s of adaptation with a particular view (shown is the front-view adaptation condition). The monkeys performed a similar task except that they were rewarded at the end of trial (**E**). **F**, In the no-adaptation condition of the electrophysiology part of the experiment, the monkeys were passively fixating on a sequence of the set of face views mixed with other faces and objects. Reward was delivered every 5 s if the monkeys could keep their gaze within the fixation window. Unlike the psychophysics part in which test stimuli were 14 images ranging from profile to front views, the presented stimulus set in the electrophysiology part consisted of 20 views covering all  $360^\circ$  of in-depth rotation. **G**, The same set of 20 face views was also used in adaptation conditions; however, they were preceded by 5 s of adaptation (front-adaptation condition is shown). The reward was delivered only at the end of the trial.

$>45^\circ$  or  $<45^\circ$  groups. Face-view stimuli were  $5^\circ$  grayscale images generated by Poser software (Smith Micro Software). An ambient light was used for illumination of Poser workspace to reduce the effect of illumination cues on perception of face-view directions. Response selection was made by the monkeys' saccade to one of the two different response points presented on the computer monitor. Response points were  $1^\circ$  circles with different colors,  $8^\circ$  to the right and left of the fixation point on the horizontal meridian. The monkey had to saccade to one of the response points to categorize the test stimuli as  $>45^\circ$  or  $<45^\circ$ . The association of target category to saccade points was reversed between the two monkeys. Categorization training was terminated when the monkeys' performance was  $>85\%$  over 500 trials. Approximately 40 trials per stimulus were collected for each condition per each subject. We also asked seven human subjects (three female, four male) to perform the same view categorization similar to monkey psychophysics. Human subjects delivered their response by pressing either of two arrow keys on a computer keyboard (right arrow key for  $>45^\circ$  category and left arrow key for  $<45^\circ$ ) (Fig. 1B). Subjects were trained to categorize the images using a feedback sound indicating their correct and wrong answers (performance  $>90\%$ ). After a training stage, subjects were asked to perform the same categorization task with or without adaptation (Fig. 1C,D). Adaptation was achieved by presentation of adapter stimuli for 350 ms, followed by 150

ms of a blank screen (Fig. 1D). This sequence was repeated for 10 times, resulting in 5 s of adaptation. Adapter stimulus was either  $0^\circ$  (profile-adaptation condition) or  $90^\circ$  (front-adaptation condition) views. At the end of each of 5 s adaptation blocks, there was a 250 ms blank period, followed by one of the 14 test stimuli, and subjects were required to categorize the stimulus similar to the no-adaptation condition. The monkeys performed a similar task (Fig. 1E). Approximately 40 trials per stimulus were collected for each condition per each subject.

Blocks of no-adaptation, front-adaptation, and profile-adaptation were presented separately. Thus, after testing all stimuli in profile-adaptation, the front-adaptation condition was performed. Trials in which the monkey broke its fixation before the test stage were considered as aborted trials and were not included in the analysis. The monkeys aborted  $\sim 13\%$  of trials on average (monkey 1: profile-adaptation, 17.32%; monkey 2: profile-adaptation, 9.32%; monkey 1: front adaptation, 15.84%; monkey 2: front adaptation, 10.02%). Note that, because adaptation was performed in blocks, aborting a trial could not change the direction of adaptation effect in the upcoming trials. Moreover, considering long durations of adaptation in each block ( $\sim 20$  min and more), the extra adaptation effect induced by aborted trails (on top of finished trails) is negligible compared with the overall adaptation effect.

To quantify the adaptation effect, we fit proportion of  $>45^\circ$  responses for the set of 14 images to the following logistic psychometric function:

$$\text{proportion of } > 45^\circ \text{ responses} = \frac{1}{1 + e^{a \times (\text{PES-test stimulus view direction})}}$$

where PES stands for point of equal selection, which is the stimulus that subject would equally report as  $<45^\circ$  or  $>45^\circ$ . Factor  $a$  determines the slope of the logistic function.

All aspects and conditions were the same for the monkey and human psychophysics part of the study except that the monkeys' eye position was monitored, whereas for human subjects eye movements were not monitored. Moreover, monkey, but not human, subjects were rewarded for correct categorization of objects in adaptation and no-adaptation conditions. Receiving the reward feedback could provide a compensatory mechanism for monkeys: because adaptation shifts the monkeys' response from a perfect categorization of stimuli toward a biased categorization away from the adapter, it can result in a reduction in the total amount of received reward. Thus, the monkeys could develop a strategy to change their categorization against the adaptation-induced bias to increase the amount of reward. Such a bias, if present, could decrease the repulsive shift magnitudes reported in this study. The monkeys performed the categorization test for both adaptation and no-adaptation conditions (in  $<1$  week) right after categorization training (which took 3–4 weeks for each monkey). The relatively shorter duration of the test compared with training phase can reduce the chance of developing a compensatory strategy. We did not find any difference between the effects of adaptation at the first half versus the second half of trials. However, the degree to which possible compensatory behavior has reduced the magnitude of adaptation effect is not measurable in our study.

## Electrophysiology setup and procedures

### Experimental conditions

**No-adaptation condition.** To determine the view-tuning characteristics of IT neurons, we recorded their responses to a set of face-view images (Fig. 1F). The stimulus set consisted of 20 views of an artificial face ranging from  $0^\circ$  (left profile) to  $360^\circ$  by  $18^\circ$  steps (the same face identity as used in the psychophysics part of the study). Stimuli were  $5^\circ \times 5^\circ$  and were presented at the center of the monitor while the monkey was holding its fixation in a  $2^\circ \times 2^\circ$  window at the center of the image. Each stimulus was presented for 250 ms with a 250 ms interstimulus interval. Twenty face views were presented together with 154 other images (25 human faces, 25 animal faces, 104 objects and object parts) in a pseudorandom order. For each neuron, we recorded 9–12 trials for each stimulus (median of 10).

**Front-adaptation conditions.** Similar to the no-adaptation condition, face-view images were presented in a pseudorandom order, but presentation of each of the 20 face-view images was preceded by 5 s of front adaptation (Fig. 1G). In this condition, there was no other image intermingled between the face views. Front-adaptation trials consisted of presentation of the front-view image for 350 ms, followed by 150 ms of blank screen afterward. This sequence was repeated for 10 times, resulting in 5 s of adaptation. Notice that, unlike the blank period between adapters (150 ms), the blank period before the test stimulus was 250 ms (Fig. 1G). For each neuron, we recorded seven to nine trials per each stimulus in this condition (median of 8).

**Profile-adaptation conditions.** It was similar to the front-adaptation condition except that the adapter was left-profile view of the same face identity. For each neuron, we recorded seven to nine trials per each stimulus in this condition (median of 8).

### Electrophysiology

Head posts and recording chambers were implanted on the skulls of two adult male macaque monkeys (*Macaque mulatta*). The same monkeys were used in the psychophysics and recording part of our experiments. Spike recordings in one of the monkeys were made before the psychophysics study and one after that. Both monkeys showed a perceptual repulsive shift in the psychophysics and neural adaptation effects in terms of response suppression and shift in neurometric function. All experimental procedures conformed to National Institutes of Health

*Guidelines on the Care and Use of Laboratory Animals.* In each recording session, a tungsten electrode (FHC) was advanced with an Evart 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 determined during electrode advancement. Additional technical details about monkey surgery and recording procedure have been described previously (Kiani et al., 2005). We recorded the responses of neurons in the lower bank of the superior temporal sulcus (STS) and the IT (extending from the lower bank of the STS toward the medial bank of the anterior middle temporal sulcus). Recordings were performed from 13 to 21 mm anterior to the ear line with 1 mm track intervals in both monkeys. Monkeys fixated on a  $0.5^\circ$  circular fixation spot presented at the center of the display in a  $2^\circ \times 2^\circ$  window. The eye position was monitored by an infrared system (i\_rec; <http://staff.aist.go.jp/k.matsuda/eye/>). The same head-fixed eye-monitoring system was used for the monkey psychophysics part of the study.

Responses of 153 IT neurons to a set of 20 face views evenly spaced between  $-180^\circ$  to  $180^\circ$  view directions were recorded before adaptation. Of 153 neurons recorded in the no-adaptation condition, 131 neurons, which showed significant responses to at least one of the face-view images, were selected for additional test under adaptation conditions.

### Analyses

The Wilcoxon's signed-rank test was used for hypothesis tests in this study unless mentioned otherwise in each instance. One-way ANOVA was used to calculate whether responses of each neuron could discriminate face views or not. The response window of 70–320 ms was used for evoked firing rate analyses.

Based on the responses of neurons to face and non-face images, we also quantified the face category selectivity defined by  $d'$  index.  $d'$  values were computed using the mean responses to each stimulus and the variance of the mean responses to each of the different stimuli as follows:

$$d' = \frac{[\mu_f - \mu_{nf}]}{\sqrt{(\delta_f^2 + \delta_{nf}^2)/2}}$$

where  $\mu_f$  and  $\mu_{nf}$  are the average responses to face and non-face images, respectively, and  $\delta_f^2$  and  $\delta_{nf}^2$  are the variance of those responses. Higher values of this index indicate a higher selectivity for faces compared with non-face images. The range of  $d'$  in the population of 131 face-responsive cells was  $-0.50$  to  $2.67$ , with mean  $\pm$  SD of  $0.22 \pm 0.58$ . The  $d'$  index was calculated using 50 human and animal faces versus 104 non-face images. The 20 face-view images were not included in the calculation of the  $d'$  index.

Normalization was done separately for non-adaptation and each of the adaptation conditions. The reason for separate normalization was the effect of adaptation on baseline responses. However, the main findings reported in Figures 4–8 were present when responses were normalized together (across all conditions) with or without removing the baseline. Note that Figure 3 illustrates a different normalization for illustration purposes. To denote the baseline activity of neurons, Figure 3 illustrates the normalized firing rates for different conditions when neuronal responses to 20 presented views are normalized between 0 and their maximal response in each condition.

The responsive range was defined as the proportion of face views driving the response of a neuron significantly above its baseline ( $-100$  to  $0$  ms from stimulus onset) firing rate ( $p < 0.05$ , Kolmogorov–Smirnov test). Values ranged from 0 to 1. Responsive range is defined based on nine trials for each stimulus (minimum number of trials across all neurons). For determining the  $F$  value in ANOVA analysis (view sensitivity), we also used nine trials.

The latency measurements in Figure 8 are performed over a population of neurons. For each population of  $N$  neurons, the average normalized visual response, the response change, and the shift magnitude are measured in each temporal bin of 50 ms starting 100 ms before the stimulus onset until 250 ms after. The normalized visual response is

quantified by averaging the responses to all 20 presented stimuli in the no-adaptation condition keeping the time course of responses. Then, dividing this average response by the overall response of the neuron to all stimuli across time normalizes the average response of this neuron. For each population of  $N$  neurons in each time point (Fig. 8A, gray rectangle), we average the responses of the  $N$  neurons for that exact time. To detect the latency of responses for each population of neurons, we compare the responses between 50 ms bins. The first 50 ms bins are defined as baseline. The latency for each population is defined as the first 50 ms bin in which the values exceed significantly ( $p < 0.05$ ) above the baseline bin using a Wilcoxon's rank-sum test between the 50 values of test bin (50 ms) and the 50 values of baseline bin. Notice that, using this method, we are not smoothing the responses and also we are not including the variability across the neurons in each population. Responses for each population of  $N$  neurons are reduced to a single array of temporal responses, and deviation of those responses from its baseline (the first 50 ms bin) is considered to be the latency of neuronal responses. A similar method was used for detecting the first time of deflection in response change and shift magnitude. The reason to consider the responses as a single array was that, in the case of our analysis in Figure 8, each row consists of a different number of neurons. Considering the responses as a single average normalized response, we can detect the first time the responses take off from the baseline while ignoring how many neurons have contributed to that response. We also quantified the latency for Figure 8A–C using a method incorporating the number of neurons: averaging across time for 50 ms bins and comparing the sliding bins. Notice that, in this case, each bin includes  $n$  numbers, where  $n$  is the number of neurons in each group. Using this method (which is partly affected by the number of neurons), we could still observe the reported findings in Figure 8, namely no correlation between the latency of visual responses and the responsive range (similar to Fig. 8A) and a significantly negative correlation between the responsive range and the response change (similar to Fig. 8B) and shift magnitude (similar to Fig. 8C).

#### Support vector machine classifier

We used a support vector machine (SVM) algorithm to classify the presented face views based on responses of IT neurons. We designed different classifiers for different groups of neurons with various responsive ranges. For each group, we take a random subset of 15 neurons (to have an equal number of neurons in each group), then for each selected set of neurons, we randomly take 20% of trials out and train the classifier with remaining 80% and then test with the unseen 20% of trials. We performed the same algorithm 50 times for each group of neurons to quantify the mean and variance of the classifier performance for that group.

#### Optimal decoding method

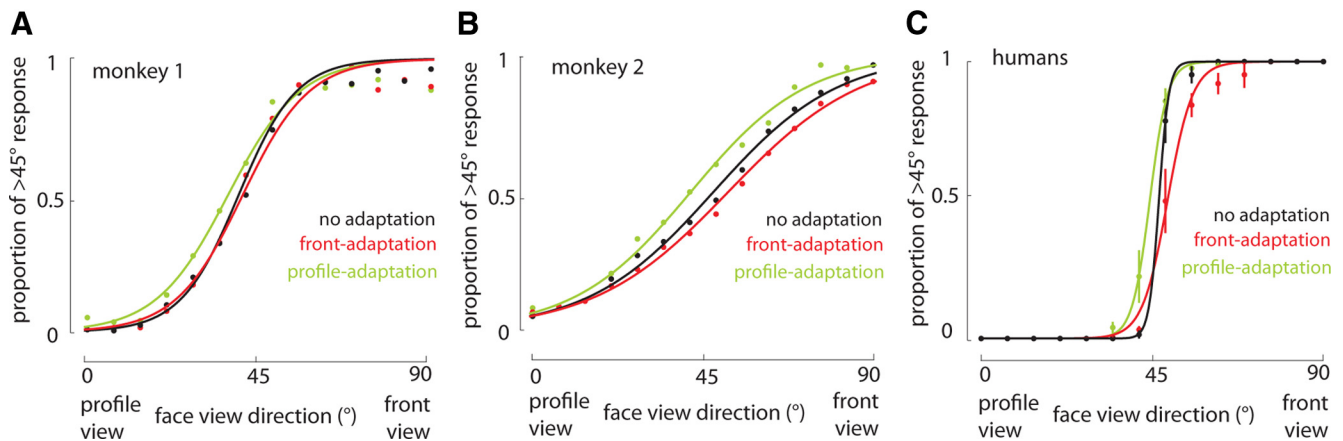
First, we created a transformation matrix based on the responses of recorded neurons (see Fig. 5). The matrix is constructed based on average adaptation-induced neuronal response change for different stimuli with various viewing direction and for neurons with various responses to the adapter in the no-adaptation condition. To construct this matrix, we used only the responses of 113 neurons responding to more than one face view. We also smoothed the resulting matrix along both dimensions (view direction and adapter response) with averaging neighboring values for each point (e.g., view direction of  $120^\circ$  and adapter response of 0.5 in the final matrix is the average of values for view directions between  $100^\circ$  and  $140^\circ$  and adapters of 0.45 to 0.55). This smoothing helped reduce jitters in responses of model units. These two fine adjustments (i.e., removing minimally responsive neurons and smoothing with neighboring values) helped to remove the noise in the resulting matrix. As shown in the transformation matrix in Figure 5, whereas maximal adaptation-induced suppression (blue shading) is around where the absolute relative distance between the tested stimulus and preferred direction was small and adapter was optimal (top and middle), this suppression reduced for stimuli farther from the adapter (moving farther from 0 on abscissa) or for non-optimal adaptations (moving away from top toward bottom on ordinate). Moreover, the response enhancement (red shading) was greatest in the case of non-optimal adaptation (bottom part of the matrix).

At the next step, we developed a dense model of tuned units and estimated the predicted response of model units after adaptation using their transformation matrix. Basically, for each neuron, by knowing the adapter optimality (the response before adaptation) and the response to each stimulus, we could assign a value in the transformation matrix. Thus, we could predict the expected change attributable to adaptation for each stimulus for different neurons with different responses to the adapter. We used an optimal decoding method (Dayan and Abbott, 2001) based on responses of model units before and after adaptation to decode the responses of the units of the model before and after adaptation. In this method, by knowing the responses of  $N$  units to  $S$  stimuli, we can predict the most probable presented stimulus given a set of  $N$  responses (say responses of  $N$  units to one stimulus after adaptation). Thus, for a set of responses to each stimulus before and after adaptation, we can decode the most probable presented stimulus. In other words, it can reveal that the given response set (after adaptation) looks more similar to responses of which stimuli before adaptation. We performed the same algorithm to responses of model units to different stimuli before and after adaptation with 0 and  $90^\circ$  adapters to see whether the decoded stimulus is  $>45^\circ$  or not. Repeating the same algorithm 100 times with added Poisson noise (independently added to responses before and after adaptation), the probability of decoding each stimulus  $>45^\circ$  was calculated. Thus, we could quantify the psychometric function of the model of probability of reporting each stimulus as  $>45^\circ$  (similar to humans' and monkeys' function). PES was measured for each psychometric function before and after adaptation. To compare PESs before and after adaptation, we performed the same assessment 100 times, giving us 100 estimates of PES per each adaptation and no-adaptation conditions.

## Results

### Presence of face-view aftereffect in humans and monkeys

To examine the neural signature of face-view aftereffect, we first verified its presence in monkey subjects with a set of parametric face views. Two macaque monkeys performed a view categorization task with no adaptation or with 5 s of adaptation to either front or profile views (Fig. 1A,E). Based on the subject's responses, the stimulus that the subject categorized equally as similar to front and profile views was defined as the PES. We observed evidence of a repulsive perceptual shift in the monkeys' responses after adaptation (Fig. 2A,B). The PES was significantly decreased after adaptation with profile view [monkey 1:  $\Delta\text{PES}_{\text{profile}} (\text{PES}_{\text{profile-adaptation}} - \text{PES}_{\text{no-adaptation}}) = -2.78$ ; monkey 2:  $\Delta\text{PES}_{\text{profile}} = -5.98$ ;  $p < 0.01$  for both, logistic regression]. A significant increase in the PES was observed after front-view adaptation in both monkeys, indicating a lower probability of categorizing the stimuli into the  $>45^\circ$  group (monkey 1:  $\Delta\text{PES}_{\text{front}} = 1.06$ ; monkey 2:  $\Delta\text{PES}_{\text{front}} = 3.97$ ;  $p < 0.01$  for both, logistic regression). Significant PES enhancement after front adaptation and reduction after profile-adaptation both indicate the presence of a repulsive perceptual shift. To show that similar view aftereffect is also present in humans, we asked seven subjects to perform a similar view categorization task. Consistent with previous reports (Fang and He, 2005), after adaptation to the front view, human subjects also demonstrated a decrease in probability of categorizing images into the  $>45^\circ$  group, whereas profile-adaptation increased this probability ( $\Delta\text{PES}_{\text{front}} = 2.57 \pm 0.90$ ,  $p = 0.025$ ;  $\Delta\text{PES}_{\text{profile}} = -2.58 \pm 0.87$ ,  $p = 0.029$ ) (Fig. 2C). The differences between the monkey and human psychophysics paradigms (e.g., the presence of reward feedback and eye monitoring for monkeys, which were absent for humans), in addition to differences in subjects' sensitivity (humans were more sensitive as indicated by their sharper psychometric function) makes it hard to compare the magnitude of repulsive shift between humans and monkeys. However, the presence of a perceptual repulsive shift in



**Figure 2.** *A, B*, Repulsive perceptual shift away from the adapter in responses of two monkey subjects for stimuli from 0° to 90°. The y-axis indicates the probability of categorizing test stimuli (x-axis) as >45° group. The rightward shift of psychometric curves after front-adaptation and their leftward shift after profile-adaptation compared with no adaptation indicates the repulsive perceptual shift attributable to adaptation. The same trend was also observed in seven human subjects performing the same task (*C*). Error bars indicate 1 SE.

both humans and monkeys indicates the presence of view aftereffect across species.

### Responses of IT neurons to face-view adaptation

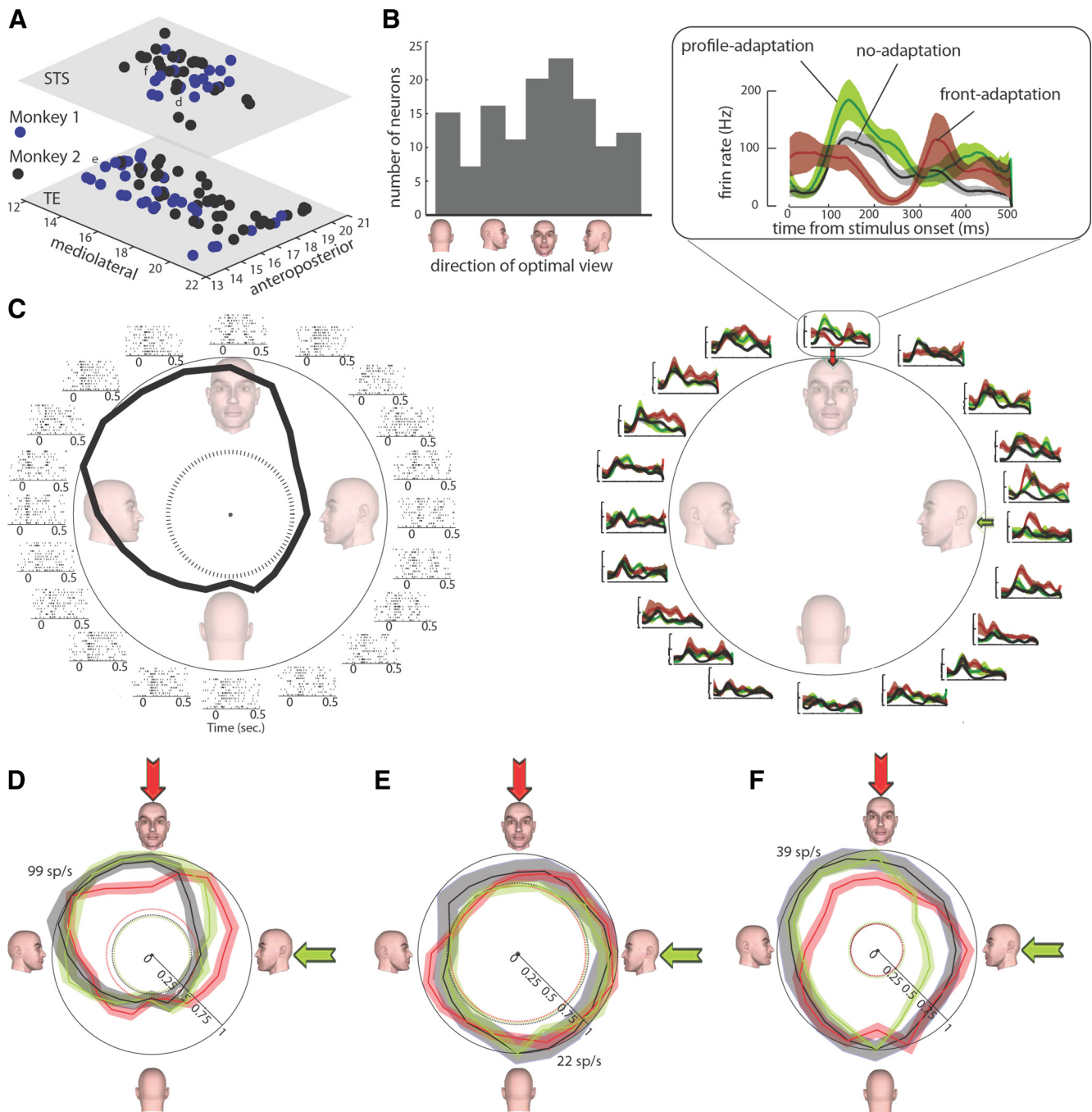
To study the neuronal correlates of the observed view aftereffect, we recorded responses of 131 neurons in the IT (the lower bank of the STS and the temporal cortex) of two passively fixating monkeys before and after view adaptation. Figure 3, *A* and *B*, shows the coordinates and response profiles of the recorded neurons.

Figure 3*C* illustrates responses of a view-sensitive IT neuron. The rastergram shows responses to 20 presented face views before adaptation. This neuron exhibits its highest response to views between front and right-profile views. The polar plot illustrates the average response of this neuron for each stimulus in a time window of 250 ms starting 70 ms after the stimulus onset. The dashed circle indicates the baseline firing rate. This neuron was responsive to 50% of the presented face-view stimuli (responsive range, 0.5; peak/baseline firing rate, 2.52). The effect of adaptation on this neuron is shown on the right, in which the black trace and the gray shading indicate the mean  $\pm$  SE of responses before adaptation, and the green and the red shading indicate the responses in profile and front-adaptation conditions. Separately normalized average responses (to maximal stimulus response) of this neuron to all views in adaptation (red and green for front and profile-adaptation, respectively) and no-adaptation (black) conditions are shown in Figure 3*D*. Response to the adapter decreased from 0.94 to 0.68 and from 0.49 to 0.47, respectively, in front- and profile-adaptation conditions. Spiking activity of a neuron responding to a narrower range of face views is illustrated in Figure 3*E*. This neuron was responsive to 25% of the face-view range (responsive range, 0.25). Its peak firing rate was 1.4 times greater than its baseline firing rate. The response to front view was 0.81 before adaptation, which reduced to 0.78 after front adaptation. Profile adaptation reduced the response to adapter from 0.93 to 0.91. The wide responsive range neuron illustrated in Figure 3*F* was responsive to the whole face-view range (responsive range, 1; peak/baseline firing rate, 3.78). The response to front view was 0.90 before adaptation, which reduced to 0.72 after front adaptation. Profile-adaptation reduced the response to adapter from 0.79 to 0.51.

We found a significant decrease in baseline activity of IT neurons after adaptation. The average baseline activity was  $18.40 \pm 1.24$  Hz before adaptation, which reduced to  $13.45 \pm 0.94$  Hz

after front adaptation and to  $14.24 \pm 0.94$  Hz after profile adaptation ( $p < 10^{-3}$  for both adaptation conditions). To dissociate the effects of adaptation on baseline responses from those on visual responses, the visual responses are normalized separately before and after the adaptation condition between maximum and minimum visual response. Consistent with these three exemplar neurons, we found adaptation-induced suppression in the responses of most IT neurons. We found that the normalized responses of neurons to the adapter were significantly decreased after adaptation [ $\Delta\text{response}_{\text{front-adapter}} = -0.10 \pm 0.02$  (mean  $\pm$  SE),  $p < 10^{-3}$ ;  $\Delta\text{response}_{\text{profile-adapter}} = -0.08 \pm 0.02$ ,  $p < 10^{-3}$ ]. Considering the preference of the neuron for the two adapters before adaptation, we found this response suppression to be greater for the preferred adapter (two-way ANOVA: adaptation,  $p < 10^{-3}$ ; adapter preference,  $p < 10^{-3}$ ; interaction,  $p = 0.002$ ).

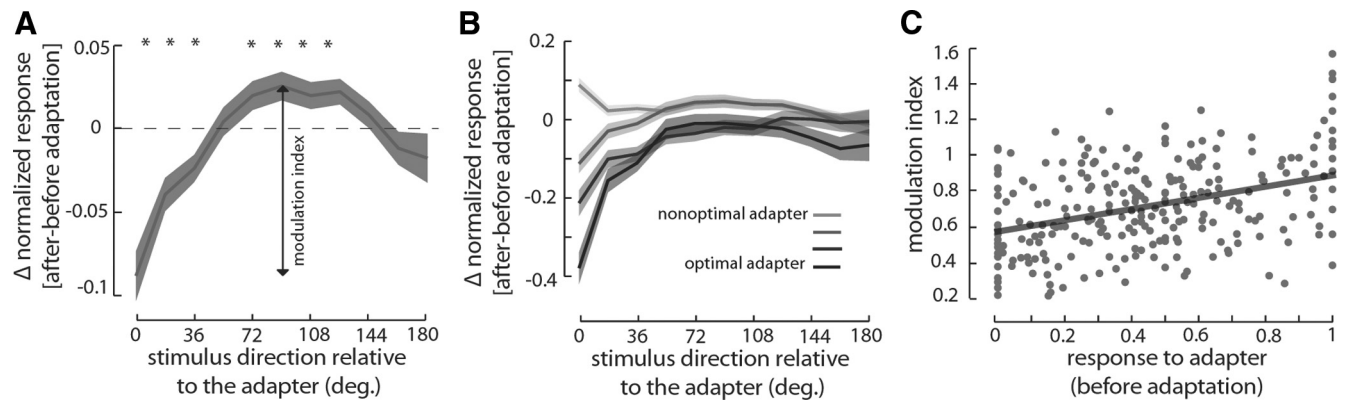
Moreover, we found that the adaptation-induced response suppression spreads to other stimuli that share features with the adapter. Figure 4*A* illustrates the average response change for stimuli with various view direction offsets relative to the adapter. The adaptation-induced response suppression gradually decreased for stimuli farther away from the adapter in view direction, leading to adaptation-induced response enhancement for stimuli with  $\sim 50^\circ$  or more distance from the adapter. The magnitude of the adaptation-induced response change also depended on the initial response of the adaptor. To quantify the effect of adaptation, we defined the modulation index as the difference between maximum enhancement and maximum suppression attributable to adaptation (Fig. 4*A*). As shown in Figure 4*B*, the adapter suppression is greater when adapter is an optimal stimulus, whereas adaptation with non-optimal adapters results in response enhancement rather than suppression, suggesting that optimal adapters modulate the responses more than non-optimal adapters. Across neurons, the initial response to the presentation of the adapter positively correlated with the modulation index (Pearson's correlation = 0.37;  $p < 10^{-3}$ ), suggesting that the effect of adaptation might be generated within the recorded population (Fig. 4*C*). To summarize, adaptation generally resulted in response suppression for the adapter and stimuli with close views. However, we observed adaptation-induced response enhancement in two cases: (1) when responses are tested for stimuli farther than  $\sim 70^\circ$  away from the adapter view (Fig. 4*A*); and (2) when the adapter is a non-optimal stimulus (Fig. 4*B*, palest trace). These enhancements suggest the presence of intercon-



**Figure 3.** *A*, Anteroposterior and mediolateral coordinates of 131 recorded neurons in temporal cortex (bottom surface) and STS (top surface) of two monkeys (blue and black). The coordinates of three exemplar neurons shown in *D*, *E*, and *F* are indicated in this panel by letters *d*, *e*, and *f*, respectively. *B*, Histogram illustrating the distribution of optimal view for 131 recorded neurons. *C*, Responses of an exemplar STS neuron responsive to 50% of the view range before and after adaptation. Left, Rastergram of responses to 20 presented face views in the no-adaptation condition. The dashed circle indicates the baseline firing rate. Right, The polar plot illustrates the average response of this neuron for each stimulus in no-adaptation (black), profile-adaptation (green), and front-adaptation (red) conditions (as indicated in the inset). The shading indicates the SE of responses in each condition. Response of this neuron to all views in adaptation (red and green for front- and profile-adaptation, respectively) and no-adaptation (black) conditions is shown in *D*. Responses are normalized separately, and the maximum response of 99 spikes/s (sp/s) in the no-adaptation condition is indicated. The dashed lines denote the baseline response in each condition, and the shading indicates the SE of responses in each condition. *E* and *F* illustrate responses of two other neurons responsive to 25 and 100% of the view range, respectively, with the same conventions as in *d*. Note that, in *C–F*, responses are separately normalized between zero and peak response (which is different from normalizations used to report the findings in Results) only for illustration purposes of showing the baseline activity.

nected and competing neural assemblies that are specifically tuned to various face views with antagonistic interactions. In this realm, adaptation of the cells tuned to the adapter view suppresses their spiking activities and consequently, presumably via disinhibition, enhances activities of other cell groups that respond non-optimally/have tunings to views far away from the adapter stimulus.

To fully study the changes of IT responses after adaptation, we quantified the same measures with different normalization methods. When responses were normalized across conditions, we still observed adaptation-induced response suppression for the adapter stimulus [ $\Delta\text{response}_{\text{front-adapter}} = -0.07 \pm 0.02$  (mean  $\pm$  SE),  $p < 10^{-3}$ ;  $\Delta\text{response}_{\text{profile-adapter}} = -0.06 \pm 0.02$ ,  $p < 10^{-3}$ ]. We also observed the significant positive relationship between the modula-



**Figure 4.** The effect of adaptation on responses of IT neurons. **A**, Adaptation decreases responses to the adapted view and views close to it, whereas responses to stimuli farther from the adapter are enhanced. The y-axis indicates the mean and SE of change in response attributable to adaptation for stimuli with various offset in viewing direction relative to the adapter (x-axis). The modulation index is defined as difference between maximal enhancement and suppression after adaptation. **B**, The effects of different strengths of adapter on responses of IT neurons. The darkest trace indicates the responses when adapter was optimal ( $>0.75$ ), whereas the palest trace shows those responses when the adapter is non-optimal ( $<0.25$ ). The two other traces are adaptation-induced responses for adapters between 0.25 and 0.5 and adapters between 0.5 and 0.75. **C**, The effect of adaptation is a factor of adapter strength. The x-axis indicates the adapter strength, i.e., the normalized response to the adapter before adaptation. The y-axis indicates the modulation index. The black line depicts the regression line fitting the data (modulation index =  $0.577 + 0.313 \times$  adapter response).

tion index and the adapter strength (Pearson’s correlation = 0.22;  $p < 10^{-3}$ ). The same three significant effects were observed with other normalization methods: (1) subtraction of baseline response and then normalizing separately for different conditions ( $p < 10^{-3}$  for all three effects); and (2) subtraction of baseline response and then normalizing all conditions together ( $p < 10^{-3}$  for all three effects). Also, when we quantified the adaptation effect as the ratio of responses before and after adaptation (after-adaptation response/before-adaptation response), this measure was significantly smaller than 1, indicating a response suppression attributable to adaptation (response-ratio<sub>front-adapter</sub> =  $0.53 \pm 0.07$ ,  $p < 10^{-3}$ ; response-ratio<sub>profile-adapter</sub> =  $0.64 \pm 0.08$ ,  $p < 10^{-3}$ ). Using this measure of adaptation effect, we also observed a significant positive relationship between the modulation index and the adapter strength (Pearson’s correlation = 0.31;  $p < 10^{-3}$ ). Moreover, the observed increase in response to stimuli farther from the adapter was independent of our method of quantification. Similar to what is shown in Figure 4A, we observed significant response enhancements for stimuli farther from the adapter when the adaptation effect was quantified by a ratio of response or as a response difference after baseline subtraction. Thus, independent of the method of quantification (after/before or after – before) or the method of normalization (within each condition without baseline removal or across conditions after baseline removal), we observed three major effects of adaptation: (1) a response suppression for the adapter; (2) a response enhancement for stimuli farther from the adapter; and (3) a positive relationship between modulation of responses attributable to adaptation and the strength of adapter itself.

**Adaptation-induced response changes in IT neurons can account for view aftereffect**

It has been shown previously that IT neurons respond to face images (Gross et al., 1972) and they are sensitive to the in-depth rotation of faces (Perrett et al., 1992). However, as has been shown previously, being responsive to a stimulus or a particular dimension does not necessarily imply a contribution to the perceptually congruent representation of that stimulus or dimension (Logothetis and Schall, 1989; Kohn and Movshon, 2004). For example, Kohn and Movshon showed that, whereas both V1 and V5 neurons respond to visual motion and change their response attributable to motion adaptation, only neurons in V5 exhibit tuning function changes consistent with the observed repulsive perceptual shift in behavior (Kohn and

Movshon, 2004). To examine whether our observed effects of adaptation on IT responses could account for the psychophysically assessed repulsive perceptual shift, we tested the effects of adaptation on the relationship of stimuli in the space of neuronal responses. If IT neurons are neuronal correlates of face-view perception, we expect to see repulsion away from the adapter in the space of neuronal responses consistent with the observed repulsive perceptual shift in behavior.

We developed a model for decoding the presented stimulus based on the responses of artificial view-tuned units. The model consists of a dense population of units tuned to various view directions (Fig. 5). We predicted the adaptation-induced response change in the model units based on the observed effects of adaptation in IT neurons. Given the stimulus–response relationship of model units before adaptation, the most likely stimulus to produce a particular response after adaptation was determined using a maximum likelihood inference (Dayan and Abbott, 2001). Based on the probability of being decoded as a  $>45^\circ$  stimulus, we defined a PES for the model response in each condition. The PES was  $44.80 \pm 0.06$  before adaptation, which decreased to  $37.55 \pm 0.13$  after  $0^\circ$  adaptation ( $p < 10^{-3}$ ). Consistent with repulsive perceptual shift, the PES also significantly increased after  $90^\circ$  adaptation ( $50.57 \pm 0.19$ ,  $p < 10^{-3}$ ) (Fig. 5). These findings indicate that the direction and magnitude of the observed shift with the decoding method are consistent with those observed in the behavior of humans and monkeys in our experiment.

As an additional test, we quantified changes in the IT population response as a result of adaptation. Intuitively, we can assume the responses of the IT population at each moment as a point in a high-dimensional space whose dimensions correspond to individual neuronal responses. We investigated how adaptation moved this point. To do this, we calculated the similarity between the response to each stimulus and the response to the front view, based on the following equation:

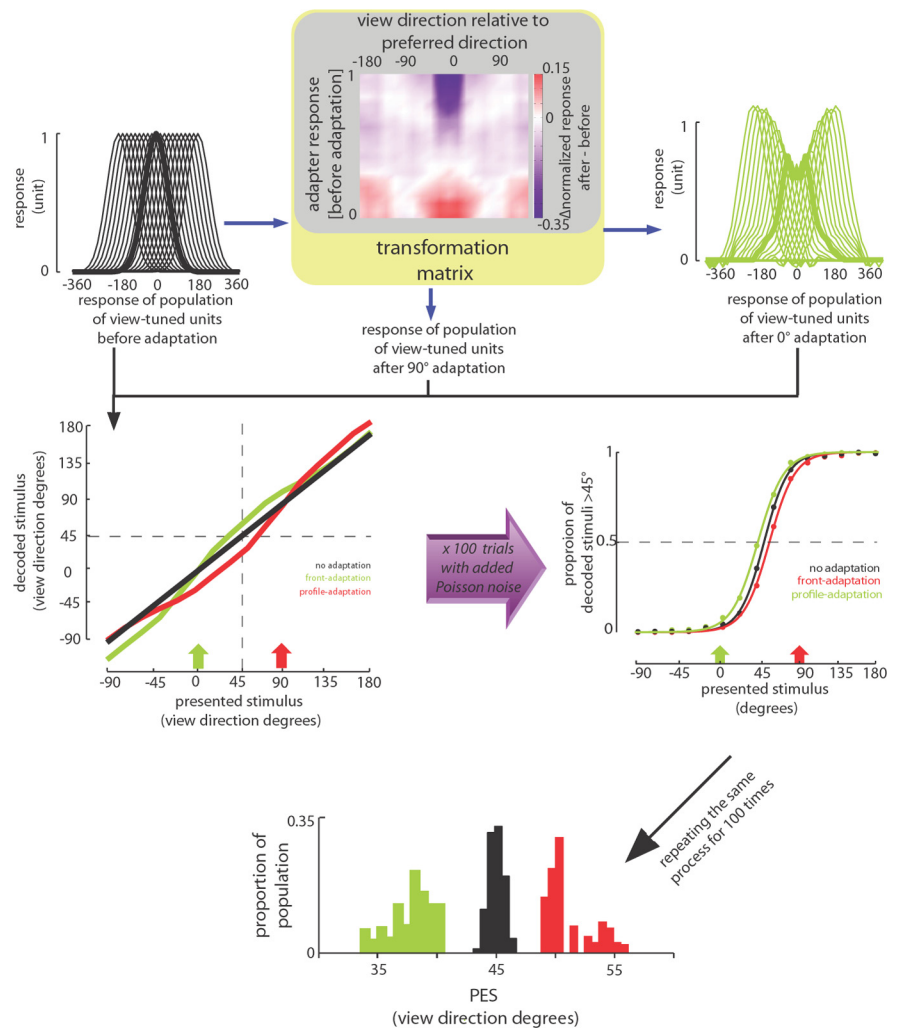
$$\text{Front similarity} = \frac{\text{Dissimilarity to left profile view}}{\text{Dissimilarity to left profile view} + \text{Dissimilarity to front view}} \times 90.$$

To quantify the dissimilarity to, for example, the left-profile view, we measured the sum of the square of the differences between the response to the test image and to the left-profile view for 131

neurons (i.e., square of Euclidian distance in the 131-dimensional space of neuronal responses). Thereby, the similarity of the test image to the front view was 90 for the front-view image and 0 for the left-profile image. The similarity to the front-view image was measured for the stimuli ranging between the left profile to the front view based on this equation. We fit a sigmoid neurometric function for the correspondence between the similarity to the front view (in space of neuronal responses) and the viewpoint direction of the presented stimuli. Using this method, we constructed the neurometric function, for no-adaptation, front-adaptation, and profile-adaptation conditions, separately (Fig. 6). We defined the “neural PES” as the stimulus that was equally distant from the profile and front stimuli. In both adaptation conditions, a repulsive shift away from the adapter was observed in neurometric functions. Neural PES before adaptation was 45.76, which increased to 54.87 after profile adaptation and decreased to 37.35 after front adaptation. Shift magnitude was defined as the change in the neural PES between profile-adaptation and front-adaptation conditions ( $54.87 - 37.35 = 17.52$ ). To quantify the magnitude of the repulsive shift, we derived samples of 130 of 131 recorded neurons based on a leave-one-out jackknife strategy. Using all possible permutations, the shift magnitude was measured for each sample and was corrected using the jackknife correction. The shift magnitude after jackknife resampling was  $18.03 \pm 4.10$ , which was significantly  $>0$  ( $p < 10^{-3}$ ). Thus, the findings with both neurometric and decoding methods are consistent with each other.

### Contribution of different classes of neurons in representing face views

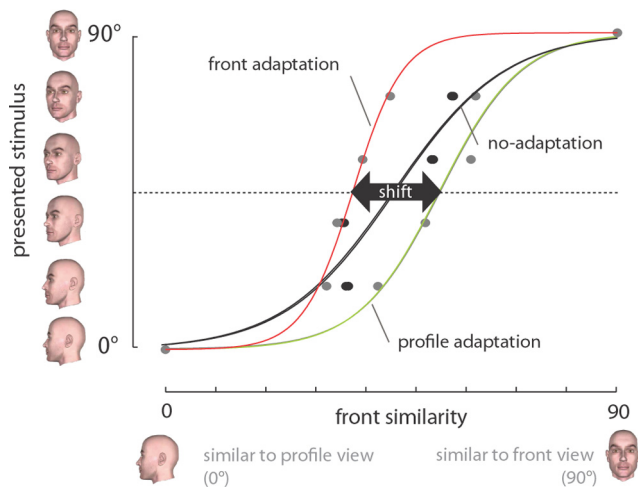
Our findings using neurometric and decoding methods both indicate that IT neurons change their responses after adaptation congruent to viewpoint aftereffect, suggesting their role as a neuronal correlate of face-view perception. Unlike the decoding, the neurometric method gives us the possibility to assess the shift magnitude for subsets of recorded neurons. To explain the neuronal correlates of view representation, we investigated the contribution of a subpopulation of IT neurons in perceptual repulsive shift. We characterized the functional class of neurons using an approach similar to previous studies (Perrett et al., 1992; Logothetis and Pauls, 1995). The “responsive range” for each neuron was defined as the proportion of face-view images that drive its response significantly above the baseline-firing rate. Similar to previous studies, we found a spectrum of neurons responding across the whole breadth of tuning, from a small range to the whole span of view changes. We found a significant positive correlation between the re-



**Figure 5.** The optimal decoding method. The transformation matrix (the color map at top) constructed based on response changes of neurons recorded before and after adaptation is shown at top. It depicts the average adaptation-induced change of response to various stimuli (abscissa) for neurons with various responses to the adapter (ordinate). Responses of model units before adaptation are shown at top left. Responses of units after  $0^\circ$  adaptation are predicted using the transformation matrix and are shown at top right. Decoded stimuli based on response of model units to various stimuli are plotted at middle left. Because the decoding is based on the responses before adaptation, the decoded and presented stimuli follow the unity line before adaptation (black line). After adaptation with  $0^\circ$  stimulus (green arrow), the decoded stimuli show a repulsive shift away from the adaptation point (green line). For quantification, we added independent Poisson noise to responses and decoded them. This process was repeated 100 times. The probability of being decoded as  $>45^\circ$  is shown at middle right for different stimuli and different conditions. Based on this probability, we defined the PES in each condition as the stimulus that would be decoded as  $>45^\circ$  in half of the trials. The bottom shows distribution of PESs before adaptation (black), after adaptation with  $90^\circ$  adapter (red), and  $0^\circ$  adapter (green) for 100 iterations of the above-mentioned process.

sponsive range of neurons and their ability to discriminate faces versus non-face objects, as quantified by a  $d'$  index (Pearson's correlation = 0.372,  $p < 10^{-3}$ ). Dividing the view-responsive range into four quartiles, we found a significant difference between these groups in their ability to discriminate between face and non-face objects ( $F = 7.94$ ,  $p < 10^{-3}$ , one-way ANOVA) (Fig. 7A). This is consistent with previous studies that found neurons with wide responsive ranges to be more face selective (Perrett and Oram, 1993; Booth and Rolls, 1998). However, a concern in our study is that a narrow-tuned neuron could have a lower  $d'$  than a view-tolerant neuron because it responds to a narrower range of face-view stimuli; and thus this could induce a relationship between  $d'$  and the face-view selectivity. We cannot quantify the degree to which the observed correlation between responsive range and face selectiv-





**Figure 6.** The neurometric method. The black line plots similarity to front view for various stimuli ranging between left profile to front view based on responses of population of 131 neurons in the no-adaptation condition. The similarity measured after profile and front adaptation is shown in green and red. “Shift” magnitude is defined as the change in front similarity of half-profile (45°) stimulus between profile- and front-adaptation conditions.

ity is induced by the similarities between the two measures. However, the stimuli we used to measure the face selectivity were mostly (>80%) front views  $\pm 45^\circ$ , which reduces the confounding effect of view direction on the face selectivity measure, at least for neurons with a wide responsive range.

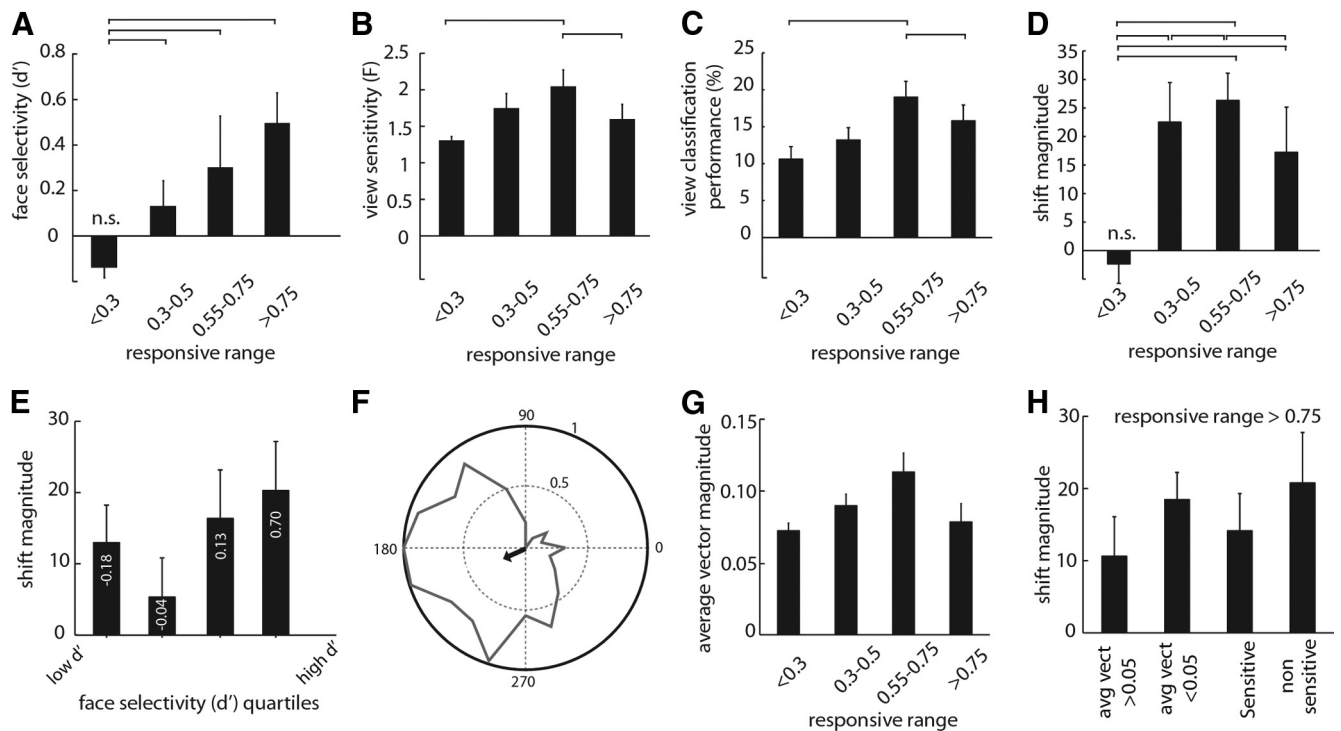
To quantify the ability of neurons to discriminate between different views, we defined the view sensitivity of each neuron as an  $F$  value of a one-way ANOVA performed over its responses to 20 face views (significant view sensitivity is defined as  $p < 0.05$ ). Face-view sensitivity was significantly different between groups of neurons with different responsive ranges ( $F = 4.15$ ,  $p = 0.008$ , one-way ANOVA) (Fig. 7B). View sensitivity was maximal in groups of neurons responding to 50–75% of face views. However, we found that neurons responding to a wide range of face views (>75% of the range) are still able to significantly discriminate between various views (view sensitivity =  $1.621 \pm 0.221$ ,  $p < 10^{-3}$ ). To verify the ability to read out the face views at the level of population of neurons, we designed a classifier based on SVM algorithms. Similar to what we observed for single-neuron-level analysis in Figure 7B, the population-level analysis also shows the difference between the ability of different groups of neurons with various responsive ranges in discriminating the face views as shown in Figure 7C ( $F = 12.36$ ,  $p < 10^{-3}$ ). Moreover, it shows that the population of neurons with large responsive ranges (>0.75) discriminate face views significantly above the chance level (performance =  $15.80 \pm 1.13$ ,  $p < 10^{-3}$ ). Thereby, both single-neuron- and population-level analyses indicate that the ubiquitous responses of these neurons to wide range of views do not interfere with their ability to discriminate different views. As shown in Figure 7D, we calculated the shift magnitude for different groups of neurons with various responsive ranges. The shift magnitude was greatest for the population of neurons responding to 50–75% of the view range (shift magnitude =  $26.33 \pm 4.77$ ). However, the population of neurons responding to >75% of views still show significant amount of shift magnitude (shift magnitude =  $17.23 \pm 7.91$ ,  $p < 10^{-3}$ ). Therefore, wide responsive range neurons not only exhibit view sensitivity in their visual response, they also modulate their response to account for perceptual repulsive shift attributable to adaptation. We also examined the relationship between the face selectivity

( $d'$ ) and the contribution in view representation (Fig. 7E). We split the neurons into four quartiles according to their face selectivity. Whereas the greatest shift magnitude was observed in neurons with the highest face selectivity, we found a significant amount of shift in responses of all four quartiles across the whole range of face selectivity ( $p < 0.05$ ). This finding implies a contribution of both face-selective and non-selective neurons in the representation of face views.

Whereas the  $F$  value in a one-way ANOVA is a good measure of the ability of a neuron to discriminate between views, it does not directly address whether the neuron is coding a particular view or not (i.e., a response to a specific stimulus). We quantified the ability of each neuron for coding particular views by calculating the average vector of the neuron, in which the normalized response to each view direction determines the vector length for that particular direction (Fig. 7F). This quantity is basically a measure of view sensitivity assuming a unimodal distribution, which ranges from 0 to 1. Higher values indicate that the responses of the neurons were consistently evoked by a particular stimulus and other stimuli similar to that, whereas low values indicate that the neuronal responses were similar across views. The mean  $\pm$  SD of the average vector for the population of recorded neurons was  $0.08 \pm 0.04$ . As expected, we found a significant correlation between the magnitude of the average vector and sensitivity (measured as  $F$  of ANOVA) ( $r = 0.59$ ,  $p < 10^{-3}$ , Pearson’s correlation). The relationship between the responsive range and the average vector magnitude (shown in Fig. 7G) was similar to that seen between the responsive range and sensitivity (Fig. 7B). We also quantified stimulus selectivity as (best – worst)/(best + worst), where best is the highest response to the set of 20 stimuli, and worst is the lowest response. Stimulus selectivity was also highly correlated with response sensitivity ( $r = 0.49$ ,  $p < 10^{-3}$ , Pearson’s correlation). Moreover, similar to view sensitivity (Fig. 7B) and the average vector method (Fig. 7G), we observed the same relationship between stimulus selectivity and responsive range ( $F = 5.25$ ,  $p = 0.001$ , one-way ANOVA).

The findings that (1) wide responsive neurons are able to signal view directions (Fig. 7B,C,G) and (2) they change their responses consistent with perceptual repulsive shift after adaptation suggest that responding to a wide range of views does not interfere with their ability to represent face views. Importantly, we found that responding to wide range of views is sufficient to play a role in view aftereffect, independent of the ability to signal a specific view direction. As shown in Figure 7H, we measured the shift magnitude for a population of wide responsive range neurons with and without significant view sensitivity. Even the small subset of non-view-sensitive wide responsive range neurons ( $n = 6$ ) exhibited a significant amount of shift in their neurometric function (shift =  $20.77 \pm 6.97$ ,  $p = 0.021$ ). We did not find a significant difference between the amount of shift between view-sensitive and non-sensitive wide responsive neurons ( $p = 0.693$ ). The same trend was observed when we looked at wide responsive neurons with selective (average vector >0.05) and non-selective (average vector <0.05) behavior (shift<sub>non-selective</sub> =  $18.46 \pm 3.75$ ,  $p = 0.033$ ; shift<sub>selective</sub> vs shift<sub>non-selective</sub>,  $p = 0.377$ ) (Fig. 7H).

We found that wide-tuned neurons are able to discriminate face views and follow adaptation-induced perceptual changes. To answer the question of whether or not the representation in wide-tuned neurons is dependent on that of narrow-tuned ones, we studied the temporal characteristics of responses and the signatures of view aftereffect for different groups of neurons. We did not find any correlation between the response onset latency of neurons and their responsive range (Pearson’s correlation =



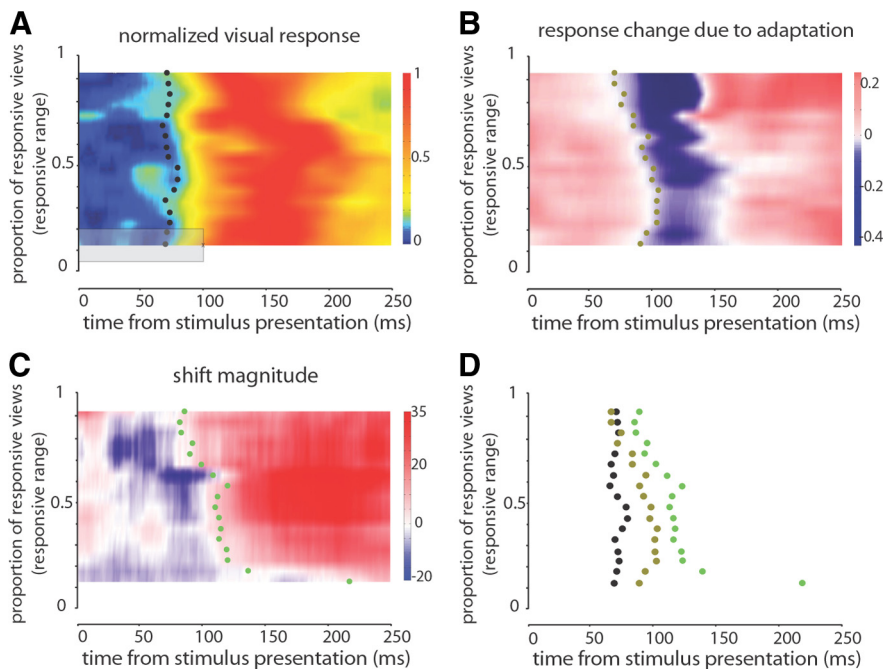
**Figure 7.** The relationship of responsive range with face-category selectivity and face-view sensitivity. **A**, The average face selectivity for different groups of neurons with various responsive ranges. **B**, The average face-view sensitivity for groups of neurons. **C**, The average performance for classifiers constructed based on responses of equal number of neurons in different groups with various responsive ranges. **D**, The average shift magnitude based on neurometric method for different groups of neurons. The lines above columns indicate the one-by-one comparisons that were significantly different from each other. All columns were significantly different from zero except those denoted by n.s. **E**, The shift magnitude for groups of neurons with different face selectivity ( $d'$ ). The  $d'$  was divided into four quartiles. The average  $d'$  for each group is denoted, and the bars indicate the shift magnitude for that group of neurons. **F**, The average vector method. Normalized responses of a sample neuron to 20 views are shown in the polar plot. The response to each stimulus is assumed as a vector in which the length of the vector is defined by the evoked response and the angle is defined by the view direction of the presented stimulus. The arrow indicates the average vector (over 20 stimuli) of this neuron, which is used as another measure of stimulus sensitivity assuming a unimodal distribution of responses. **G**, The mean and SE of the average vector magnitude for neurons with different responsive range. **H**, The shift magnitude measured for subpopulations of wide responsive range neurons (responsive range > 0.75). Shift is measured for wide-range neurons with average vector magnitude > 0.05 (1st column) or < 0.05 (2nd column), view-sensitive neurons (i.e., significant ANOVA test; 3rd column), and non-sensitive neurons (4th column).

$-0.120$ ,  $p = 0.646$ ) (Fig. 8A). However, the suppressive effect of adaptation appeared earlier in neurons responding to a wider range of views (Pearson's correlation =  $-0.819$ ,  $p < 10^{-3}$ ) (Fig. 8B). Furthermore, the signature of repulsive perceptual shift (i.e., the positive shift magnitude in the neurometric function) also appeared earlier in wide responsive range neurons (Pearson's correlation =  $-0.76$ ,  $p < 10^{-3}$ ) (Fig. 8C). For example, the average visual latency was 72 and 73 ms, respectively, for neurons responding to >80% of face views and for neurons responding to 40–55% of views. However, the wider responsive group exhibited the response suppression 24 ms earlier, and the repulsive shift in neurometric function was 26 ms earlier than the other group. Onset latency of neural responses, adaptation effects, and shift magnitudes are depicted in Figure 8D for comparison.

To examine the possibility that IT neurons receive the adaptation effects from neurons with smaller receptive fields in the lower visual areas, we recorded responses of 20 (of 131) IT neurons in a spatially variable adaptation condition. In this condition, the location of the adapter was randomly changed between corners of an  $8^\circ \times 8^\circ$  window to bypass the low-level adaptation effects. Considering the large receptive field of IT neurons (Desimone and Gross, 1979), adapters presented at peripheral locations could still induce direct response adaptations in IT neurons. Changing the adapter location weakens the adaptation effect for neurons with smaller receptive fields in the lower areas that provide feedforward visual information to the IT, thereby the spatially variable adaptation condition is more indicative for

adaptation effects that are not inherited from areas upstream in the ventral visual pathway. Nine cells were tested in the profile-adaptation condition, whereas adaptation occurred at the same location as the target (foveal profile adaptation, corresponding to our original profile-adaptation condition) and at different locations (spatially variable profile adaptation). Eleven cells were tested with front-view adapter in both foveal and spatially variable adaptation conditions. All these neurons were also tested under the no-adaptation condition. The PES was significantly decreased after foveal and remote profile adaptation compared with the no-adaptation condition (profile adaptation:  $\Delta\text{PES}_{\text{foveal}} = -8.34 \pm 1.78$ ,  $p = 0.007$ ;  $\Delta\text{PES}_{\text{remote}} = -10.79 \pm 1.25$ ,  $p = 0.003$ ). The PES in remote adaptation was not significantly different from that in foveal adaptation ( $p = 0.164$ ). Consistent with our previous findings, the PES was significantly increased after foveal and spatially variable front-adaptation compared with the no-adaptation condition (front adaptation:  $\Delta\text{PES}_{\text{foveal}} = 12.16 \pm 0.66$ ,  $p < 10^{-3}$ ;  $\Delta\text{PES}_{\text{remote}} = 15.52 \pm 0.46$ ,  $p < 10^{-3}$ ). In fact, the PES in spatially variable adaptation was significantly greater than that in foveal adaptation ( $p = 0.011$ ).

Adaptation was executed in blocks, and the order of profile and front adaptation blocks was alternated between sessions. We verified that the order of the blocks does not change the adaptation effects. Adaptation-induced response suppression for the very adapter was not significantly different between sessions when profile adapter or front adapter was presented first or second ( $p = 0.232$  and  $p = 0.803$ , respectively). Moreover, whereas



**Figure 8.** Adaptation modulates activity of neurons responding to a wide range of face views before neurons responding to a smaller range. The average normalized responses for groups of neurons with various responsive ranges are shown in **A**. To determine the value for each point (e.g., the small  $x$ ), the responses over a window of 100 ms and a responsive range of 0.3 were averaged together (indicated by the gray rectangle). The small  $x$  was shifted with steps of 1 ms and 0.05 of responsive range. Small black dots indicate the average visual latency for each group of neurons. Whereas visual latency is not a factor of responsive range, adaptation-induced response changes appear earlier in neurons with wider responsive range as it is shown in **B**. Using the same averaging window as **A**, response change, defined as change in normalized response after adaptation compared with before adaptation, is illustrated in **B**. The small dots indicate the first significant negative response change for each group of neurons. **C** illustrates the shift magnitude, the measure of repulsive shift, for different groups of neurons. The averaging window is similar to **A** and **B**. Note that the first significant positivity in shift magnitude appears earlier for neurons with wide responsive range. Small dots indicate the first significant positive shift for each group of neurons. **D** summarizes the relative timing of visual latency, onset of response suppression, and onset of positive change in shift magnitude for various groups of neurons.

the no-adaptation condition was always presented before the adaptation conditions, after adaptation we tested the responses of IT neurons one more time to check whether the responses “recovered” after adaptation. Thus, the recovery block only consisted of presentation of view stimuli without any adaptation. Similar to the comparison between adaptation and no-adaptation conditions, we found an adaptation-induced response suppression (similar to Fig. 4) and a repulsive shift of neurometric functions (similar to Fig. 6) when we compared adaptation with the recovery condition. Thus, IT neurons could recover their response after adaptation.

## Discussion

We found an adaptation-induced perceptual repulsive shift in human and monkeys. This finding supports the idea that, in the primate brain, there are specialized units that process head-view direction. We further provided evidence for the specific role of neurons in the IT of monkeys that account for the observed perceptual repulsive shift, suggesting that the activity of these neurons could serve as the neural basis of face-view representation. Moreover, contrary to common expectation from modeling studies, we found that neurons responding to a wide range of views are able to discriminate between face views and surprisingly exhibit correlates of view aftereffect before neurons responding more selectively to narrower range of views. Here we discuss the three main implications of this finding.

First, our finding implies a level of visual processing that encodes the presence of an object despite its pose-dependent variations (wide responsive range) while keeping its sensitivity to those variations (significant view sensitivity; Fig. 7*B,C*). Neurons at this stage show the highest face selectivity (Fig. 7*A*). Interestingly, concomitant with their wide-range responses to face views, they also convey face-view information (Fig. 7*B,C*). These findings argue against the necessity of a tradeoff between object selectivity and metric invariance and show the presence of shape-selective neurons with wide, but graded, view responses in IT (Rust and Dicarlo, 2010). We have shown previously the presence of neurons in IT that respond selectively to object categories (e.g., animate body) but also discriminate subgroups of the selected category (Kiani et al. 2007). Generalization across a particular stimulus dimension (such as face views or object category) while discriminating the constituent subgroups suggest concurrent processing of multiple levels of visual information by IT neurons.

Second, our findings have implications for the interpretation of psychophysical and functional brain imaging experiments that use adaptations and aftereffects. For example, Fang and He (2005) showed that the view aftereffect exists in human subjects, and this was interpreted as psychophysical evidence for the existence of view-selective neurons and thereby the existence of a viewer-centered

representation in humans (Culham, 2005; Fang and He, 2005). Here we studied the responses of single neurons in IT and, in contrast to the interpretation of the findings of Fang and He, found that wide-range neurons may have a role in this aftereffect. Similarly, the assumption that wide responsive range neurons do not play any role in view-dependent object recognition is also foundational in the interpretation of some adaptation fMRI experiments. In many fMRI adaptation paradigms, the spread of response suppression to non-adapted views after adaptation to a specific view is interpreted as a measure of neuronal response tuning in that area (Grill-Spector et al., 1999; Grill-Spector and Malach, 2001; Andrews and Ewbank, 2004). However, as indicated by our single-neuron data (Figs. 3*F*, 7*D*, 8*C*), view specific, adaptation-induced modulation of responses is observed in neurons responding to a limited or a large range of views. This evidence emphasizes the concern raised previously by other groups about the dissociation between selectivity and adaptation effects (Sawamura et al., 2006). According to our findings, whereas the presence of face-view aftereffect can potentially imply the presence of view sensitivity, one needs to be cautious about interpreting it as evidence for the presence of view-tuned units.

Last, our results imply that, whereas neurons that are believed to form a view-specific representation (narrow-range tuning) are involved in the representation of face views, those responding to the presence of faces regardless of view point (wide-range tuning) also presumably play a role in the representation of face views.

Interestingly, neurons in the latter group exhibit changes in their response earlier than the former group, suggesting that the view information in wide-range neurons is probably not necessarily received from view-tuned units.

It has been suggested that a view-centered representation is formed at the level of IT view-tuned neurons as a proxy for view-dependent recognition. Based on modeling studies, it is widely assumed that view-invariant neural responses are shaped in later stages of the feedforward and hierarchically organized ventral visual pathway by pooling the responses of the lower-level view-tuned neurons (Riesenhuber and Poggio, 1999). Here we provide evidence that face-selective IT neurons with wide-view tuning not only conveyed face-view information but also exhibited neural signatures of view aftereffect before neurons with narrower tuning (Fig. 8*B–D*). These findings that view-dependent representation of objects does not necessarily depend on narrow-tuned neurons is inconsistent with these hierarchical models and suggest the presence and role of parallel pathways in formation of view-selective and invariant representations. So the alternative possibility, as raised previously by some models of object recognition, is that viewpoint-dependent recognition of objects just necessitates different neural paths with various sensitivity to the in-depth rotations “en route” rather than to the existence of viewpoint-specific representations (Bar, 2001). A possible explanation for how wide-range neurons exhibit signatures of view aftereffect earlier than narrow-range neurons is that the former receive their information not from the latter group but from a faster and more direct route. However, this speculation requires more direct analysis about the connectivity of IT neurons to understand the real nature of hierarchical organization of neurons with various tuning to in-depth rotation. It is important to acknowledge here that neurons in the ventral stream can be tuned to various degrees to different features, such as in-depth rotation, tilt, or position of an object simply by being sensitive to a particular feature necessary for identification of that object. These neurons play their role in extracting those features, but the final representation of, for example, tilt or view is not essentially shaped at the level of these neurons. We suggest that the goal of visual object recognition system is to accumulate sensitivity and tolerance at its final stage, to be read out accordingly based on the task at hand.

## References

- Afraz SR, Kiani R, Esteky H (2006) Microstimulation of inferotemporal cortex influences face categorization. *Nature* 442:692–695. [CrossRef Medline](#)
- Andrews TJ, Ewbank MP (2004) Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage* 23:905–913. [CrossRef Medline](#)
- Bar M (2001) Viewpoint dependency in visual object recognition does not necessarily imply viewer-centered representation. *J Cogn Neurosci* 13:793–799. [CrossRef Medline](#)
- Booth MC, Rolls ET (1998) View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb Cortex* 8:510–523. [CrossRef Medline](#)
- Culham JC (2005) Turn the other cheek: viewpoint aftereffects for faces and objects. *Neuron* 45:644–645. [CrossRef Medline](#)
- Dayan P, Abbott L (2001) *Theoretical neuroscience*. Cambridge, MA: Massachusetts Institute of Technology.
- Desimone R, Gross CG (1979) Visual areas in the temporal cortex of the macaque. *Brain Res* 178:363–380. [CrossRef Medline](#)
- Fang F, He S (2005) Viewer-centered object representation in the human visual system revealed by viewpoint aftereffects. *Neuron* 45:793–800. [CrossRef Medline](#)
- Grill-Spector K, Malach R (2001) fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica* 107:293–321. [CrossRef Medline](#)
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24:187–203. [CrossRef Medline](#)
- Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of macaque. *J Neurophysiol* 35:96–111. [Medline](#)
- Kiani R, Esteky H, Tanaka K (2005) Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. *J Neurophysiol* 94:1587–1596. [CrossRef Medline](#)
- Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J Neurophysiol* 97:4296–4309. [CrossRef Medline](#)
- Kohn A, Movshon JA (2004) Adaptation changes the direction tuning of macaque MT neurons. *Nat Neurosci* 7:764–772. [CrossRef Medline](#)
- Logothetis NK, Pauls J (1995) Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb Cortex* 5:270–288. [CrossRef Medline](#)
- Logothetis NK, Schall JD (1989) Neuronal correlates of subjective visual perception. *Science* 245:761–763. [CrossRef Medline](#)
- Logothetis NK, Sheinberg DL (1996) Visual object recognition. *Annu Rev Neurosci* 19:577–621. [CrossRef Medline](#)
- Perrett DI, Oram MW (1993) Neurophysiology of shape processing. *Image Vision Comput* 11:317–333. [CrossRef](#)
- Perrett DI, Hietanen JK, Oram MW, Benson PJ (1992) Organization and functions of cells responsive to faces in the temporal cortex. *Philos Trans R Soc Lond B Biol Sci* 335:23–30. [CrossRef Medline](#)
- Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nat Neurosci* 2:1019–1025. [CrossRef Medline](#)
- Rust NC, Dicarlo JJ (2010) Selectivity and tolerance (“invariance”) both increase as visual information propagates from cortical area V4 to IT. *J Neurosci* 30:12978–12995. [CrossRef Medline](#)
- Sawamura H, Orban GA, Vogels R (2006) Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the fMRI adaptation paradigm. *Neuron* 49:307–318. [CrossRef Medline](#)
- Tanaka K (1996) Inferotemporal cortex and object vision. *Annu Rev Neurosci* 19:109–139. [CrossRef Medline](#)