Duration-Dependent fMRI Adaptation and Distributed Viewer-Centered Face Representation in Human Visual Cortex

Fang Fang[1], Scott O. Murray[2] and Sheng He[1]

[1]Department of Psychology, University of Minnesota, Minneapolis, MN 55455, USA and [2]Department of Psychology, University of Washington, Seattle, WA 95108, USA

Two functional magnetic resonance imaging (fMRI) face viewpoint adaptation experiments were conducted to investigate whether fMRI adaptation in high-level visual cortex depends on the duration of adaptation and how different views of a face are represented in the human visual system. We found adaptation effects in multiple face-selective areas, which suggest a distributed, viewer-centered representation of faces in the human visual system. However, the nature of the adaptation effects was dependent on the length of adaptation. With long adaptation durations, face-selective areas along the hierarchy of the visual system gradually exhibited viewpoint-tuned adaptation. As the angular difference between the adapter and test stimulus increased, the blood oxygen level–dependent (BOLD) signal evoked by the test stimulus gradually increased as a function of the amount of 3-dimensional (3D) rotation. With short adaptation durations, however, face-selective areas in the ventral pathway, including the lateral occipital cortex and right fusiform area, exhibited viewpoint-sensitive adaptation. These areas showed an increase in the BOLD signal with a 3D rotation, but this signal increase was independent of the amount of rotation. Further, the right superior temporal sulcus showed little or very weak viewpoint adaptation with short adaptation durations. Our findings suggest that long- and short-term fMRI adaptations may reflect selective properties of different neuronal mechanisms.

Keywords: adaptation, face, fMRI, viewpoint, visual cortex

Introduction

A hallmark feature of the human visual system is its ability to recognize most objects from different viewing angles. Two competing theories of object recognition propose different explanations. The viewer-centered theory (e.g., Ullman 1989; Poggio and Edelman 1990) suggests that recognition is based on matching specific views to a set of templates, which requires explicit viewer-specific object representations. The object-centered theory (e.g., Biederman 1987) suggests that it is based on constructing a structural description of simple parts, which does not require explicit representations of objects from specific views. Very recently, a viewpoint aftereffect was found to strongly support the existence of viewer-centered object representations in the human high-level visual system (Fang and He 2005). Specifically, after visual adaptation to an object viewed 30° from one side, when the same object was subsequently presented near the front view, the perceived viewing directions were biased in a direction opposite to that of the adapted viewpoint.

Although our previous behavioral adaptation experiment supported the idea that objects and faces are represented by populations of viewpoint-tuned neurons, how these representations are implemented in the visual cortex is still unknown. In

the current study, we focus on how different views of a face are represented in the human visual cortex. For example, what are the viewpoint-tuning properties of the various face-selective occipital–temporal areas? How are these properties established along the hierarchy of the visual system? Are there cortical areas in the visual system that respond to faces in a viewpoint-independent fashion? We used functional magnetic resonance imaging (fMRI) adaptation experiments to address these questions. Previous functional magnetic resonance adaptation experiments (Grill-Spector and others 1999; Andrews and Ewbank 2004) have given partial answers to the above questions. However, these studies provided limited information about specific viewpoint-tuning properties. In addition, a blocked adaptation design may have confounded adaptation with attention (Kanwisher and Yovel 2006).

Adaptation has served as a powerful psychophysical tool for demonstrating selective neural sensitivities to various stimulus dimensions, from low-level stimulus features (Kohler and Wallach 1944; Blakemore and Campbell 1969; Anstis and Moulden 1970) to high-level object and face properties (Webster and Maclin 1999; Leopold and others 2001; Zhao and Chubb 2001; Suzuki and Grabowecky 2002; Rhodes and others 2003; Watson and Clifford 2003; Webster and others 2004; Fang and He 2005). In these studies, subjects were typically exposed to adapting stimuli for long durations (tens of seconds preadaptation and several seconds topping-up adaptation, here called "long-term adaptation"), and the subsequently perceived visual distortions in the test stimuli were used to infer how visual properties are represented in the human visual system. More recently, there has been an explosion of fMRI studies using adaptation as an experimental tool to make inferences about subvoxel level neural selectivity in specific cortical regions (Grill-Spector and others 2006; Krekelberg and others 2006).

Unlike the long-term adaptation designs traditionally used in psychophysical and single-unit studies, most of the fMRI adaptation studies, especially those related to object representations in occipital–temporal cortical visual areas, have used brief (e.g., 300 ms) adaptation times (Kourtzi and Kanwisher 2000, 2001; here called "short-term adaptation"). Though the pattern of results in these fMRI studies is often consistent with the adaptation logic, there has been little validation of the technique, and we know little about its underlying mechanism. For example, orientation-selective fMRI adaptation in V1 can be easily demonstrated using long-term adaptation (Fang and others 2005; Larsson and others 2006) but not using short-term adaptation (Boynton and Finney 2003; Murray and others 2006). It has been recognized that the duration of adaptation has a strong influence on the susceptibility of an area to

adaptation, especially in the early visual cortex (Krekelberg and others 2006). Importantly, there is evidence to suggest that the temporal dynamics of high-level behavioral adaptation are similar to those of traditional low-level adaptation effects (Leopold and others 2005). For example, just as with low-level adaptations, the face identity aftereffect grows logarithmically stronger as a function of adaptation time and exponentially weaker as a function of test duration. Similarly, in our recent behavioral viewpoint adaptation experiment (Fang and He 2005), there was a large effect of adaptation duration with the perceptual effect substantially reduced with short adaptation times (200 ms). Given the critical effect of adaptation duration in V1 measured with fMRI and the importance of duration in behavioral face adaptation, it is natural to ask, in high-level visual cortex (e.g., face- and object-selective areas), how fMRI adaptation is dependent on the duration of adaptation. Surprisingly, to the best of our knowledge, there are no fMRI or single-unit studies that have explored this question. To answer the questions we raised above, we conducted 3 event-related fMRI experiments, including long-term, short-term, and no adaptation experiments for each subject. The critical difference between them was the duration of adaptation.

Materials and Methods

Subjects

A total 7 healthy subjects (4 females and 3 males) participated in all the experiments. Three of them also participated in our previous psychophysical study (Fang and He 2005). All subjects were right handed, reported normal or corrected-to-normal vision, and had no known neurological or visual disorders. Ages ranged from 25 to 35 years. They gave written informed consent in accordance with procedures and protocols approved by the human subjects review committee of the University of Minnesota.

Stimuli and Designs

To define retinotopic visual areas, subjects passively viewed 2 types of retinotopic mapping stimuli (Sereno and others 1995; Engel and others 1997). The first were counterphase flickered (10 Hz) checkerboard wedges of 9° radius located at the horizontal and vertical meridians. These served to map boundaries between visual areas. The second were foveal (0°–2°) and peripheral (2°–7°) counterphase (10 Hz) annuli that served to map the retinotopic extent of each area. Two retinotopic mapping scans were performed—one that alternated the horizontal and vertical meridian stimuli and one that alternated the foveal and peripheral ring stimuli. In both scans, the stimuli were presented in 20-s blocks with 10 alternations between conditions, each lasting 400 s. Subjects were asked to maintain their fixation at a central cross, and this instruction also applied to other fMRI experiments in this study.

A block-design scan was used to define the regions of interest (ROIs), including face-selective areas, nonface-selective areas, and responsive areas in the primary visual cortex (V1). Subjects passively viewed images of faces, nonface objects, and texture patterns (scrambled faces), which subtended 9.4° × 9.4° and were centered at the fixation. Images appeared at a rate of 2 Hz in blocks of 15 s, interleaved with 15-s blank blocks (Fig. 3). Each image was presented for 300 ms, followed by 200-ms blank interval. Each block type was repeated 5 times in the scan, which lasted 450 s.

In the long-term, short-term, and no adaptation experiments, the adapter and test stimuli were generated by projecting a 3-dimensional (3D) face model with different in-depth rotation angles (30° and 60°) onto the monitor plane, with the front view as the initial position. Both left and right rotations were used. When R60 (L60) was used as the adapter, R60 (L60), R30 (L30), and L30 (R30) were used as test stimuli. Thus, the angular differences between adapter and test stimuli were 0°, 30°, and 90°, respectively. The test stimuli were named Test0, Test30, and Test90 accordingly (Fig. 1). In the long-term and short-term adaptation experiments, the adapters (R60 or L60) were fixed in each adaptation scan, and they were balanced within subjects (R60 for 4 scans and L60 for the other 4 scans). All the adapter and test stimuli extended no more than 5.5° × 5.5°. A single 3D face model was used in these experiments.

For the long-term adaptation experiment (Fig. 2), each 400-s adaptation scan (total 8) consisted of 64 continuous trials and began with 25-s preadaptation. There were 4 types of trials—Test0, Test30, Test90, and Fixation trials. In the Test0, Test30, and Test90 trials, after 0.4-s blank interval, a test stimulus was presented for 0.3 s, followed by 0.3-s blank interval and 5-s topping-up adaptation for the next test stimulus. In the Fixation trials, an adapter was presented for 6 s during the whole trial. To minimize low-level adaptation effects, during the preadaptation, topping-up adaptation, and Fixation trials, the adapter floated randomly within a 9.4° × 9.4° area, whose center was coincident with the fixation. Its starting point was randomly distributed in this area, and its floating velocity was 0.7°/s. The positions of test stimuli were also randomly distributed within that 9.4° × 9.4° area. Subjects were asked to judge the viewing direction (left or right) of test stimulus as quickly as possible, and this instruction also applied to the short-term and no adaptation experiments.
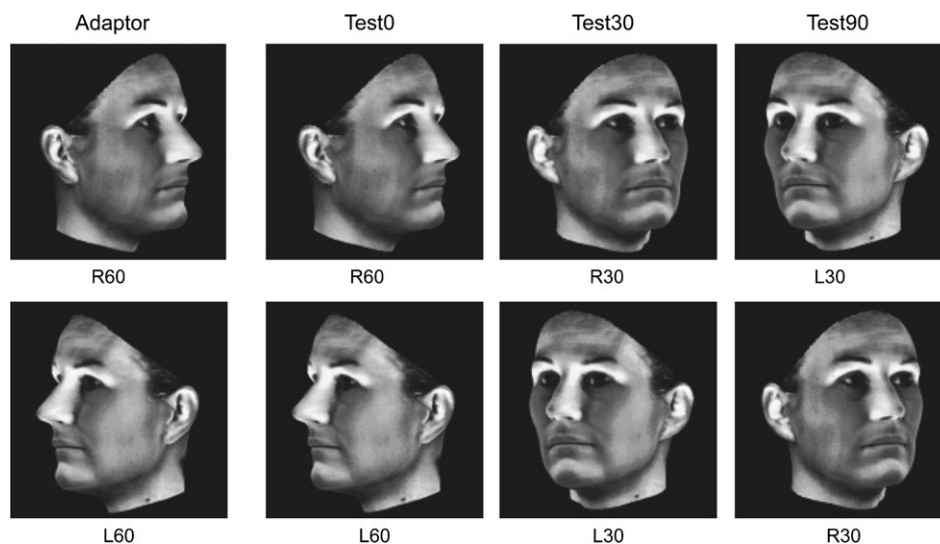

Adaptor	Test0	Test30	Test90
R60	R60	R30	L30
L60	L60	L30	R30

Figure 1. Adapters and test stimuli. Adapters are two 60° side views of a 3D face model, viewing either left or right. Test stimuli were generated by rotating the 3D face model from the adapters in-depth by 0°, 30°, and 90°, respectively, as indicated by the number following "Test."
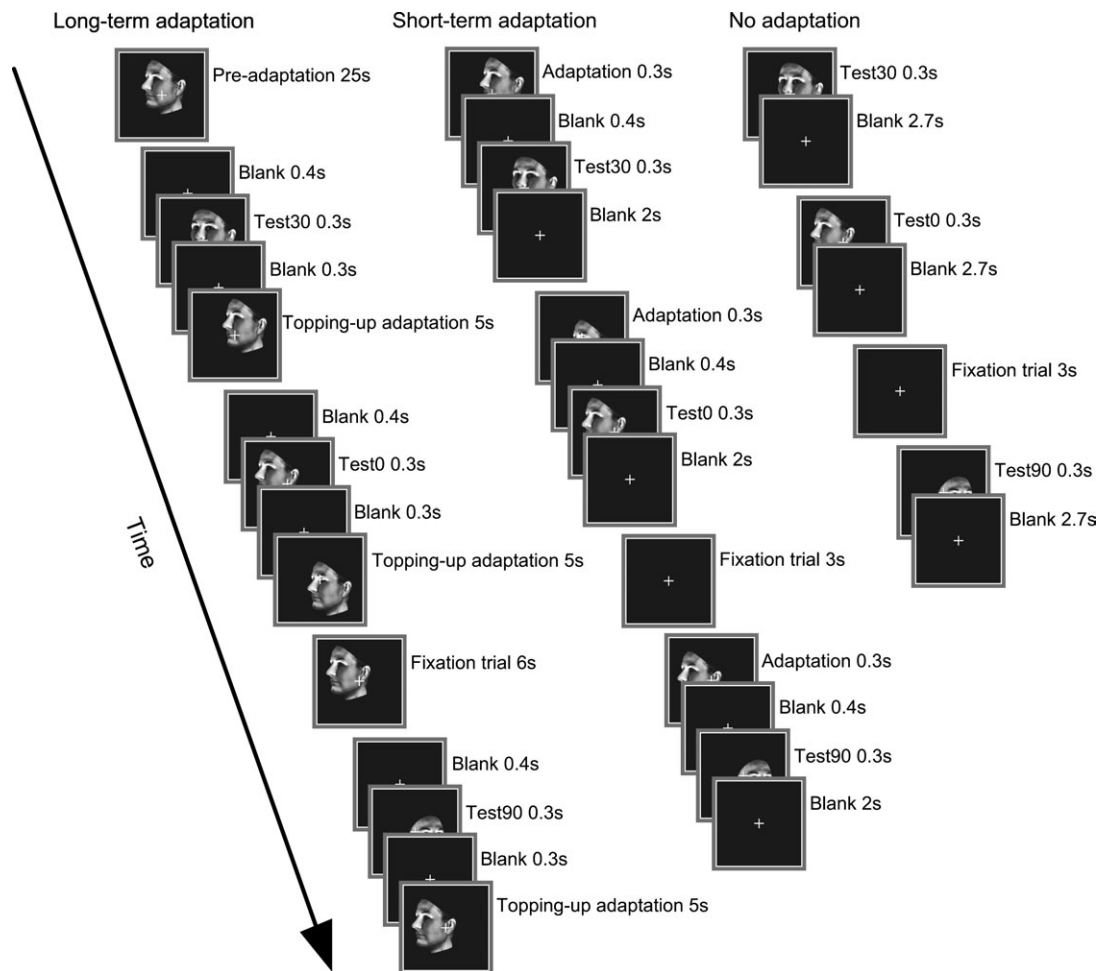
Figure 2. Schematic descriptions of the long-term adaptation, short-term adaptation, and no adaptation experiments. In the long-term adaptation experiment, test stimuli were preceded by 25-s preadaptation and 5-s topping-up adaptation. In the short-term adaptation experiment, test stimuli were preceded by only 0.3-s adaptation.

The short-term and no adaptation experiments (Fig. 2) each consisted of eight 210-s scans, and each scan consisted of 64 continuous trials. Like the long-term adaptation experiment, there were 4 types of trials—Test0, Test30, Test90, and Fixation trials. For the short-term adaptation experiment, in the TEST0, Test30, and Test90 trials, after 0.3-s adaptation and 0.4-s blank interval, a test stimulus was presented for 0.3 s, followed by 2-s blank interval. For the no adaptation experiment, in the Test0, Test30, and Test90 trials, a test stimulus was presented for 0.3 s, followed by 2.7-s blank interval. For both of these 2 experiments, in the Fixation trials, there was only a blank interval lasting 3 s. The positions of test stimuli (and the adapters in the short-term adaptation experiment) were randomly distributed within that $9.4° \times 9.4°$ area.

For each of the 3 event-related experiments, there were a total of 64×8 trials, 128 for each type of trial. The order of the 4 types of trials (Test0, Test30, Test90, and Fixation) was counterbalanced across 8 adaptation scans using M-sequences (Buracas and Boynton 2002). These are pseudorandom sequences that have the advantage of being perfectly counterbalanced n trials back (we tested up to 10 trials back), so that each type of trials was preceded and followed equally often by all types of trials, including itself.

Before the subjects were scanned, subjects were given 128 practice trials. In addition, each subject participated in a psychophysical viewpoint adaptation experiment (Fang and He 2005); all exhibited a strong viewpoint aftereffect.

Magnetic Resonance Imaging Data Acquisition

In the scanner, the stimuli were back projected via a video projector (60 Hz) onto a translucent screen placed inside the scanner bore.

Subjects viewed the stimuli through a mirror located above their eyes. The fMRI data were collected using a 3-T Siemens Trio scanner with a high-resolution 8-channel head array coil. Blood oxygen level–dependent (BOLD) signals were measured with an echo planar–imaging sequence (time echo: 30 ms, time repetition: 1000 ms, field of view: 22×22 cm^2, matrix: 64×64, flip angle: 60°, slice thickness: 5 mm, gap: 1 mm, number of slices: 14, slice orientation: axial). The bottom slice was positioned at the bottom of the temporal lobes. T_2-weighted structural images at the same slice locations and a high-resolution 3D structural data set (3D magnetization-prepared rapid gradient echo; $1 \times 1 \times 1$ mm^3 resolution) were collected in the same session before the functional runs. Each subject participated in 4 fMRI sessions. One session was conducted to define retinotopic areas, and the other 3 were dedicated to the long-term, short-term, and no adaptation experiments, respectively. These 4 sessions were conducted on different days. The temporal order of the long-term and short-term adaptation experiments was randomized across subjects, and the no adaptation experiment was run last. The block-design ROI scan was conducted in all the long-term, short-term, and no adaptation sessions.

Magnetic Resonance Imaging Data Processing and Analysis

The anatomical volume for each subject in the retinotopic mapping session was transformed into a brain space that was common for all subjects (Talairach and Tournoux 1988) and then inflated using BrainVoyager 2000. Functional volumes in all the sessions for each subject were preprocessed that included 3D motion correction using SPM99, linear trend removal, and high-pass (0.015 Hz) (Smith and others 1999) filtering using BrainVoyager 2000. The images were then aligned

to the anatomical volume in the retinotopic mapping session and transformed into Talairach space. The first 10 s of BOLD signals were discarded to minimize transient magnetic saturation effects.

A general linear model procedure was used for ROI analysis. Face-selective areas were defined as areas that respond more strongly to faces than nonface objects ($P < 10^{-4}$, uncorrected). Nonface-selective areas were defined as voxels that passed the opposite contrast. The ROI in V1 was defined as area that responds more strongly to texture patterns (scrambled faces) than blank screen ($P < 10^{-4}$, uncorrected) and confined by the V1/V2 boundaries defined by the retinotopic mapping experiment.

The event-related BOLD signals were calculated separately for each subject and each experiment, following the method used by Kourtzi and Kanwisher (2000). For each event-related scan, the time course of magnetic resonance signal intensity was first extracted by averaging the data from all the voxels within the predefined ROI. The average event-related time course was then calculated for each type of trial by selective averaging to stimulus onset and using the average signal intensity during the Fixation trials as a baseline to calculate percent signal change. Specifically, in each scan, we averaged the signal intensity across the 16 trials for each type of trial at each of 13 corresponding time points (seconds) starting from the stimulus onset. These event-related time courses of signal intensity were then converted to time courses of percent signal change for each type of trial by subtracting the corresponding value for the Fixation trial and then dividing by that value. Because M-sequences have the advantage that each type of trials was preceded and followed equally often by all types of trials, the overlapping BOLD responses due to the short interstimulus interval are removed by this averaging procedure. The resulting time course for each type of trial was then averaged across scans for each subject and then across subjects.

To more directly compare the effect of adaptation duration on the rotation-dependent responses, we calculated an index of adaptation strength by removing potential scaling effects. Following the method used by Murray and Wojciulik (2004), we normalized the peak responses to reflect the proportional increase with respect to the Test0 condition: $[(P_\theta/P0)-1]$, where P_θ are the peaks for the different rotations (Test0, Test30, or Test90) and $P0$ is the peak for the 0° rotation (Test0).

Eye Movement Recording
Eye movements of 5 subjects in the long-term, short-term, and no adaptation experiments were recorded with a long-distance optics module of ASL eyetracker (Applied Science Laboratories, Bedford, MA) in the 3-T magnet.

Results

Behavioral Responses to Test Stimuli
In the long-term, short-term, and no adaptation experiments, subjects' responses to the test stimuli were both accurate and fast. The correct rates for the 3 types of test stimuli in all experiments were above 98%. The reaction times (Mean ± standard error) were as follows: 436 ± 21 ms (Test0), 452 ± 22 ms (Test30), and 443 ± 20 ms (Test90) for the long-term adaptation experiment, 440 ± 33 ms (Test0), 460 ± 30 ms (Test30), and 458 ± 28 ms (Test90) for the short-term adaptation experiment, and 486 ± 18 ms (Test0), 493 ± 20 ms (Test30), and 500 ± 19 ms (Test90) for the no adaptation experiment. For all the experiments, there were no significant performance (both reaction time and correct rate) differences between the trial types and no significant correlation between performance and the peak values of the event-related BOLD signals.

Eye Movements
The measurements indicated that eye movements were small and gaze distributions were not systematically different across trial types. Further statistical analyses showed no significant

difference across trial types in the mean eye position (long-term adaptation—x position: $F_{2,10} = 0.677$, $P = 0.535$; y position: $F_{2,10} = 0.093$, $P = 0.913$; short-term adaptation—x position: $F_{2,10} = 0.76$, $P = 0.499$; y position: $F_{2,10} = 0.179$, $P = 0.84$; no adaptation—x position: $F_{2,10} = 0.688$, $P = 0.53$; y position: $F_{2,10} = 0.277$, $P = 0.765$), the mean saccade amplitude (long-term adaptation—x position: $F_{2,10} = 1.047$, $P = 0.395$; y position: $F_{2,10} = 0.431$, $P = 0.664$; short-term adaptation—x position: $F_{2,10} = 0.398$, $P = 0.684$; y position: $F_{2,10} = 0.706$, $P = 0.522$; no adaptation—x position: $F_{2,10} = 1.072$, $P = 0.387$; y position: $F_{2,10} = 1.254$, $P = 0.336$), and the number of saccades (long-term adaptation: $F_{2,10} = 0.135$, $P = 0.876$; short-term adaptation: $F_{2,10} = 1.65$, $P = 0.251$; no adaptation: $F_{2,10} = 0.958$, $P = 0.424$). These results suggest that it is unlikely that our results could be significantly confounded by eye movements.

Region of Interest
A block-design scan (Fig. 3) was used to define ROIs, including face-selective areas and nonface-selective areas. Subjects passively viewed images of faces, nonface objects, and texture patterns (scrambled faces). Face-selective areas were defined as areas that responded more strongly to faces than nonface objects ($P < 10^{-4}$, uncorrected). Three areas were consistently found in all subjects, including the right fusiform face area
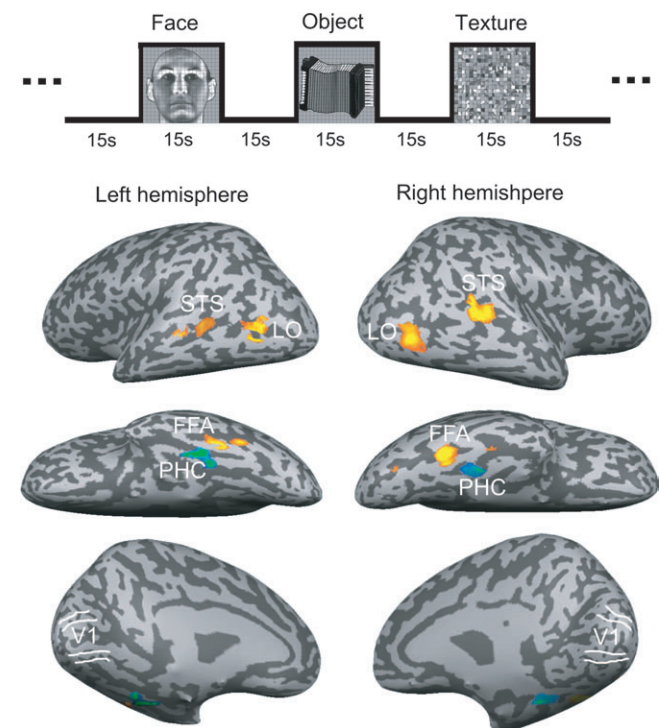

Figure 3. ROI in a representative inflated brain. The experimental procedure for ROI localization is shown at the top. Face-selective areas (red/yellow) were defined as areas that respond more strongly to faces than nonface objects ($P < 10^{-4}$, uncorrected). Blue/green areas indicate voxels that passed the opposite contrast. Face-selective ROIs consistently found in all subject include 1) rFFA (right: 33 ± 1, −40 ± 2, −14 ± 1), 2) rSTS (right: 44 ± 2, −40 ± 3, 12 ± 2), and 3) face-selective area in the LO (right: 37 ± 2, −68 ± 3, −4 ± 2; left: −41 ± 1, −65 ± 2, 2 ± 3). Nonface-selective ROIs include 1) PHC (right: 24 ± 1, −35 ± 2, −10 ± 2) rendered in blue and green and 2) primary visual cortex (V1, 0 ± 1, −84 ± 1, −2 ± 2) defined by texture stimuli and retinotopic mapping. The white lines in the bottom row are boundaries of early visual areas obtained from retinotopic mapping. The inner pair separate V1 and V2, and the outer pair separate V2 and V3.

(rFFA), the right superior temporal sulcus (rSTS), and a face-selective area in the lateral occipital cortex (LO) in both hemispheres. The face-selective LO area is also referred to as LOa (Grill-Spector and others 1999), PF (Avidan and others 2002), or OFA (Gauthier and others 2000). In addition to these face-selective areas, left superior temporal sulcus and left fusiform face area (lFFA) were found in 3 and 5 (out of 7) subjects, respectively, according to the above criterion. Non-face-selective ROIs include a region in the parahippocampal cortex (PHC) that responded more strongly to nonface objects than faces ($P < 10^{-4}$, uncorrected) and primary visual cortex (V1) defined by texture patterns and retinotopic mapping (see Materials and Methods). The PHC defined here is likely the same parahippocampal place area defined by Epstein and Kanwisher (1998), a cortical area that has been demonstrated to respond more strongly to houses and places than to other kinds of objects. For the long-term, short-term, and no adaptation experiments, this localization scan was run at the beginning of each session (total of 3 times), and the ROI locations were very similar across sessions. For fMRI data analyses, because there was no qualitative difference between the ROI locations defined in different sessions, the fMRI data presented below are from the ROIs defined in the first event-related fMRI session.

The fMRI Long-Term Adaptation Effect

After adapting to one view of a face, a cortical area that contains a collection of neural populations tuned to different views should exhibit viewpoint-tuned adaptation effect. By "tuning," we mean that, as the angular difference between the adapter and test stimulus increases, the peak amplitude of the BOLD signal evoked by the test stimulus should gradually increase as a function of the amount of 3D rotation and saturate at some angle.

A repeated-measures analysis of variance (ANOVA) of the peak amplitude revealed a significant main effect of angular difference between adapter and test stimulus in rFFA ($F_{2,14}$ = 14.581, P = 0.001), rSTS ($F_{2,14}$ = 16.028, P < 0.001), and LO ($F_{2,14}$ = 5.822, P = 0.017). In rFFA and rSTS, the BOLD signals evoked by the Test0, Test30, and Test90 showed a monotonic increase (Test90 > Test30 > Test0). This response pattern (viewpoint-tuned adaptation) was consistently observed in all 7 subjects and was confirmed by post hoc least significant difference (LSD) tests (all t > 2.6 and P < 0.04). In LO, the Test90 evoked a stronger signal than the Test30 (t = 3.022, P = 0.023). But there was no significant difference between the signals evoked by the Test0 and Test30. The nonface-selective areas V1 and right parahippocampal cortex (rPHC) did not exhibit any adaptation effect (Fig. 5).

The fMRI Short-Term Adaptation Effect

A repeated-measures ANOVA of the peak amplitude revealed a significant main effect of angular difference between adapter and test stimulus in rFFA ($F_{2,14}$ = 8.244, P = 0.006) and LO ($F_{2,14}$ = 5.934, P = 0.016) but not in rSTS ($F_{2,14}$ = 1.986, P = 0.18). We further used post hoc LSD tests to examine if there was also a monotonic increase of the peak amplitudes in rFFA and LO. It was found that the BOLD signal evoked by the Test30 was significantly stronger than that by the Test0 (both t > 2.513 and P < 0.046), but there was no significant difference between the signals evoked by the Test30 and Test90. We call this effect (Test90 = Test30 > Test0) viewpoint-sensitive adaptation effect. By "sensitive," we mean that these areas showed an increase in

the fMRI signal with a 3D rotation, but this signal increase was independent of the amount of rotation. The nonface-selective areas V1 and rPHC again did not exhibit any adaptation effect (Fig. 5).

The fMRI Responses to Test Stimuli without Adaptation

The BOLD signals in the face-selective and nonface-selective areas to different test stimuli without adaptation can be used to examine if there was any cortical response bias to a specific view of face. Repeated-measures ANOVAs showed there was no significant main effect of angular difference between adapter and test stimulus in both the face-selective areas (Fig. 4) and the nonface-selective areas (Fig. 5). We referred to the test stimuli as Test0, Test30, and Test90 here for consistency, but they essentially were faces turned 60°, 30°, and –30° from the front view. These results demonstrate that different BOLD responses to different test stimuli after adaptation are not due to a cortical response bias to a particular view of face.

Comparison between Long-Term and Short-Term Adaptation

We found viewpoint-tuned adaptation with the long-term adaptation paradigm and viewpoint-sensitive adaptation with the short-term adaptation paradigm. In order to test the effect of adaptation duration directly, we performed 2-way ANOVAs (adaptation duration [long vs. short] × test view [Test0 vs. Test30 vs. Test90]). The interaction of adaptation duration and test view was significant in rFFA ($F_{2,35}$ = 5.021, P = 0.026), rSTS ($F_{2,35}$ = 7.986, P = 0.006), and LO ($F_{2,35}$ = 9.473, P = 0.003) but not in V1 and rPHC.

To more directly compare the effect of adaptation duration between the 2 paradigms, we plotted the index of adaptation strength in Figure 6. Consistent with the statistics on the peak values, the indices were higher in the long-term adaptation experiment than that in the short-term one at the Test90 condition (at rFFA and rSTS, both t > 2.4 and P < 0.05), and there were interaction effects between adaptation duration and test view (at rFFA, rSTS, and LO, all F > 4.753 and P < 0.03).

Discussion

We found fMRI face viewpoint adaptation in multiple face-selective areas but not in the nonface-selective areas, which suggests a distributed but confined viewer-centered representation in the human visual system. This is in line with the results of Grill-Spector and others (1999) and Andrews and Ewbank (2004). More importantly, the adaptation effects were different with the long-term and short-term adaptation paradigms. Specifically, with long-term adaptation, the face-selective areas along the hierarchy of visual system gradually exhibited viewpoint-tuned adaptation. With short-term adaptation, however, face-selective areas in the ventral pathway, including LO and rFFA, exhibited viewpoint-sensitive adaptation. Another difference was that rSTS showed little adaptation effect with short adaptation durations. These differences cannot be explained simply by suggesting that longer adaptation durations lead to a stronger adaptation effect, which has been used to explain our orientation adaptation data in V1 (Fang and others 2005). Although the adaptation effect is generally stronger in rFFA and rSTS with the long-term adaptation paradigm, in LO, the short-term adaptation paradigm was even more sensitive than the long-term adaptation paradigm. Specifically, there was
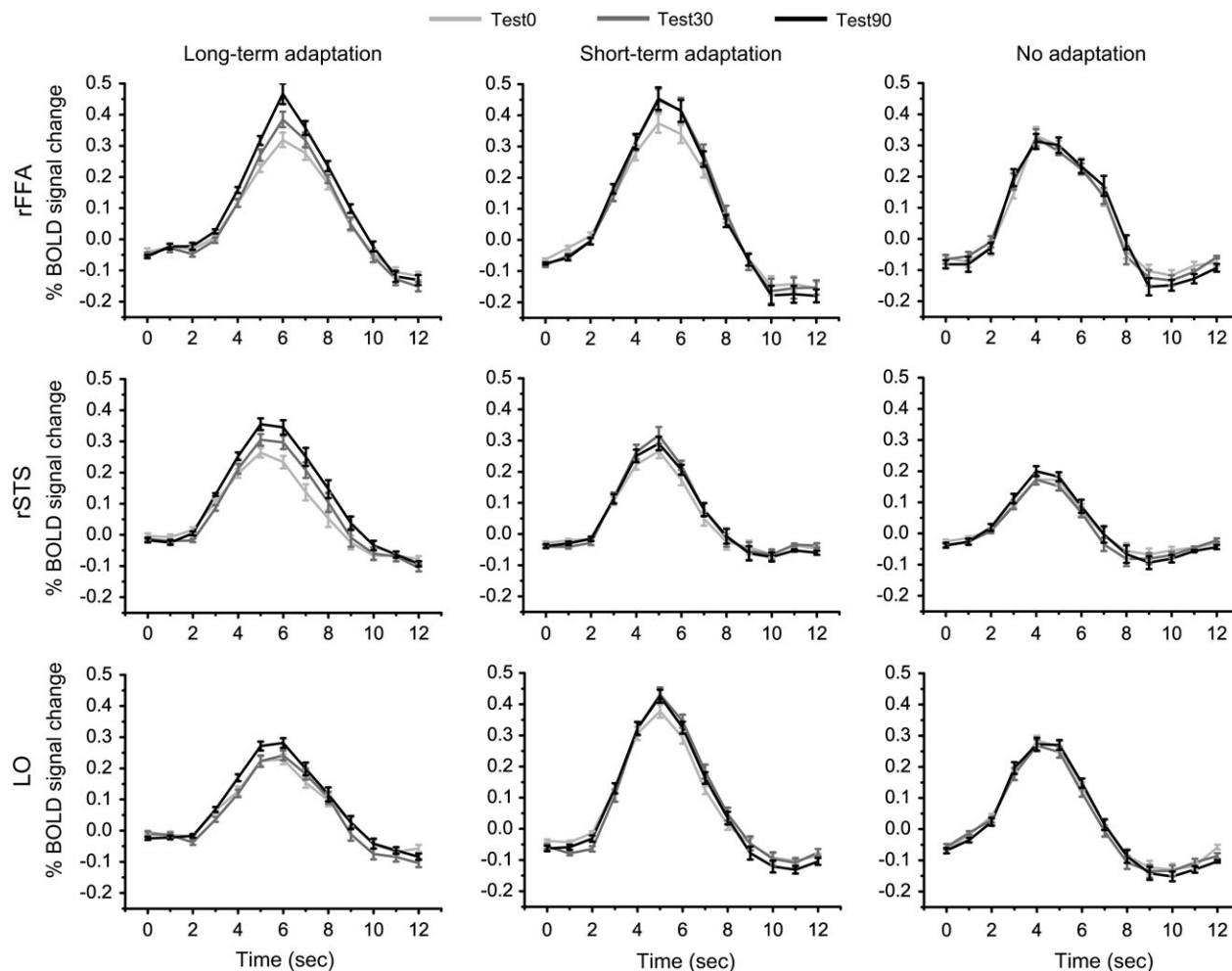
Figure 4. Time courses of event-related BOLD signals averaged across 7 subjects from the face-selective areas (rFFA, rSTS, and LO) in the long-term adaptation, short-term adaptation, and no adaptation experiments. Error bars denote 1 standard error of mean across subjects and scans at each time point.

a larger difference between the fMRI signals evoked by the Test0 and Test30 after short-term adaptation. These differences suggest that the neural mechanisms underlying the short-term and long-term adaptations may be qualitatively different—a possibility that is described in more detail below.

The no adaptation experiment was important; without it, one could argue that the observed changes in fMRI signals to different test stimuli after adaptation were due to a neural response bias to particular face views. However, in this study, there was no difference in the fMRI signal to different face views, ruling out this potential confound. In any of these 3 experiments, there was no significant performance (both reaction time and correct rate) difference between the trial types and no significant correlation between performance and the peak values of the event-related BOLD signals. Together with the null adaptation effect in the nonface-selective areas (V1 and rPHC), the overall pattern of results can be best explained by adaptation mechanisms.

Long-Term Adaptation
Neural fatigue is often suggested as a simple mechanism for adaptation effects. According to the fatigue model, all neurons initially responsive to a stimulus show a reduction in their

response after adapting to a stimulus. The reduction in a neuron's activity is proportional to its response to the adapting stimulus before adaptation. Thus, if we carefully select the test stimuli based on tuning curve properties measured in single-unit studies, we should observe viewpoint-tuned fMRI adaptation according the fatigue model. This is exactly what we found in the rSTS and rFFA in the long-term viewpoint adaptation experiment. In a previous neurophysiological study, it was found that the majority of neurons in the superior temporal sulcus (STS) (110 out of a sample of 119 responsive to the head) exhibited view selectivity, and most of them showed a unimodal tuning property (Perrett and others 1991). On average, the maximal response of these neurons was reduced by half after the head was rotated in-depth by 60° away from the optimal view. The same population of neurons was found to maintain their viewpoint selectivity irrespective of image position, size, orientation, and different lighting conditions (Perrett and others 1982, 1984, 1989). The finding of viewpoint-tuned face-selective neurons in STS has been confirmed by other groups (Desimone and others 1984; Hasselmo and others 1989; De Souza and others 2005). Such neurons were also found in the inferior temporal cortex (IT) (Desimone and others 1984). In the behavioral adaptation experiment (Fang and He 2005), we found the strength of face viewpoint adaptation effect in human
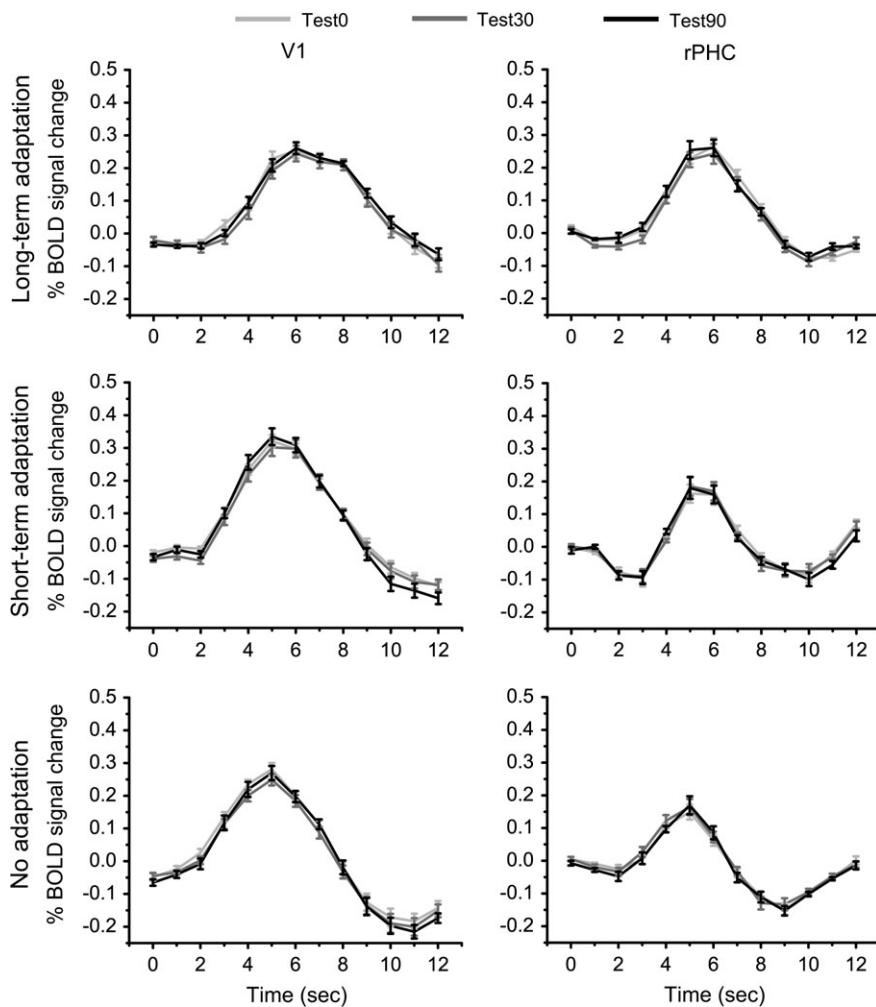
Figure 5. Time courses of event-related BOLD signals averaged across 7 subjects from the nonface-selective areas (V1 and rPHC) in the long-term, short-term, and no adaptation experiments. Error bars denote 1 standard error of mean across subjects and scans at each time point.
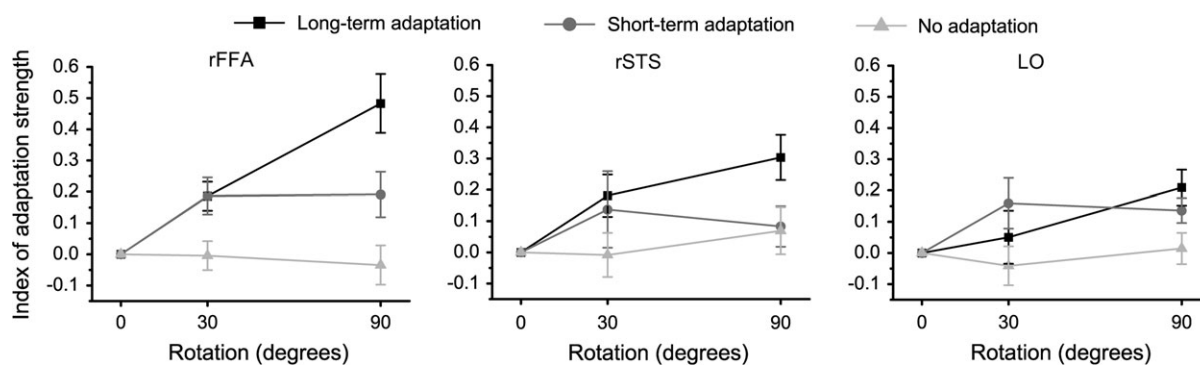

Figure 6. Indices of adaptation strength in the face-selective areas were calculated by normalizing the peak responses to the peak in the Test0 condition (see text for details). Error bars denote 1 standard error of mean across subjects and scans.

subjects depends on the angular difference between adapting and test stimuli (30° > 60°), which is parallel to the findings in these single-unit studies and the current long-term fMRI adaptation experiment.

Why was there no difference in response between the Test0 and Test30 conditions in the long-term adaptation experiment in LO? LO is generally considered to be at an early stage of

object processing, and position-invariant representations of face views have not been fully established in this area (Grill-Spector and others 1999; Kanwisher and Yovel 2006). LO may be more sensitive to the presence of specific facial parts (e.g., left or right cheek in this study). If this is the case, the lack of an adaptation effect in LO to the Test30 condition can simply be attributed to similarity of 2D image or facial part between adapters and test

stimuli, requiring a much larger rotation (90°) to evoke a differential response.

The viewpoint-tuned fMRI adaptation in rSTS and rFFA in the long-term adaptation experiment suggests a position-invariant viewer-centered face representation in these 2 areas (the positions of adapting and test stimuli were different and the adapting stimulus was slowly drifting, See Materials and Methods). The viewpoint-tuned property of rSTS and rFFA at the population level revealed with the current fMRI technique neither excludes the possibility that there are some neurons sensitive to 3D face structure in these areas (Hasselmo and others 1989; Perrett and others 1992) nor rejects the possibility that abstract view-independent face representation may exist in the memory system (e.g., medial temporal lobe) (Quian Quiroga and others 2005). Pourtois and others (2005) found viewpoint-invariant face representation in the left medial fusiform cortex using a priming paradigm and whole-brain analysis. Surprisingly, this area was clearly distinct from the face-selective areas functionally defined by a standard "face localizer," which made this result difficult to interpret. The rPHC is a scene- and house-selective area adjacent to rFFA, and it did not show face viewpoint adaptation in our study. But previous research has demonstrated viewpoint-sensitive scene adaptation in rPHC (Epstein and others 2003). Thus, the overall evidence suggests that the fMRI adaptation in high-level visual cortex is category specific.

Why do 2 cortical areas in the human visual system need to represent "redundant" information about different views of a face? The function of rFFA is thought to be face detection and/or identification (Grill-Spector and others 2004). Storing multiple views of an individual face may help the human visual system to identify a face from different viewpoints in a manner similar to viewer-centered theories of object recognition (Ullman 1989; Poggio and Edelman 1990). On the other hand, the STS is believed to be important for extracting important social interaction cues (Haxby and others 2000; Grossman and Blake 2002). If this is the case, it is not surprising that rSTS exhibited viewpoint-tuned face adaptation given the importance of head direction in conveying important information in social interaction cues.

Short-Term Adaptation

In the short-term adaptation experiment, we found viewpoint-sensitive (as opposed to viewpoint tuned) adaptation effects in rFFA and LO and little adaptation effect in rSTS. These results are consistent with previous studies—priming, repetition suppression, and rapid adaptation have been observed in LO, IT, and frontal cortex (Wagner and others 1997; Buckner and Koutstaal 1998; Buckner and others 1998; van Turennout and others 2003; Kourtzi and Grill-Spector 2005). But the short-term adaptation effect cannot be fully explained by the fatigue model based on the known single-unit and behavioral adaptation data. Desimone (1996) and Wiggs and Martin (1998) have suggested a sharpening model to explain repetition suppression effect. According to this model, some neurons that initially responded to a stimulus and coded features irrelevant to identification of the stimulus will show reduced response to subsequent presentation of that stimulus. However, the neurons that are optimally tuned to the repeating stimulus should show little or no response reductions, rather than exhibit the greatest response reductions, as in the fatigue model. As a consequence, the representation of the initial stimulus becomes sparser and

the bandwidth of tuning curves around the initial stimulus becomes narrower (see review of Grill-Spector and others 2006). This sharpening model provides a possible explanation for our viewpoint-sensitive adaptation effect in LO and rFFA. If the adapting face view had significantly narrowed the tuning bandwidth, the Test30 and Test90 would be outside the tuning range of the adapted neurons after short-term adaptation, and the subtle difference between the BOLD signals evoked by them cannot be measured by the current fMRI technique. However, if the step of transformation (e.g., rotation in depth) between adapting and test stimuli is small, the prediction of this model is that we still can observe a viewpoint-tuned adaptation effect, like the study conducted by Murray and Wojciulik (2004), which used smaller 2D rotation angles. Similar to the failure of seeing orientation adaptation in V1 with short-term adaptation, rSTS showed little or very weak effect of face viewpoint adaptation with the short-term adaptation paradigm, which suggests that the viewpoint-selective sensitivity of rSTS neurons was not significantly altered (e.g., sharpened) following brief exposure to faces and rSTS is less sensitive to the short-term adaptation than rFFA and LO. It should be noted that the short-term face viewpoint adaptation effect in rSTS may potentially be observable with fMRI—either with the development of more sensitive imaging methods or by averaging data from many more trials and subjects.

Comparing the Methods

We used 2 very different paradigms to ask the same question regarding the neural representation of face viewpoint in the human cortex. The long-term adaptation paradigm has a long tradition in psychophysics and has provided strong evidence of high-level viewer-centered neural representations (Fang and He 2005). Applying this technique in fMRI provided additional evidence about the neural correlates of these viewpoint-tuned representations in high-level visual areas, which supports the viewer-centered object recognition models. The widely used short-term adaptation paradigm has been demonstrated here again as a very sensitive tool for the same/different detection in the occipital and inferior temporal areas. Despite a substantially reduced behavioral viewpoint adaptation effect with short durations (Fang and He 2005), there is still a significant fMRI adaptation effect with short durations that strongly suggests multiple mechanisms underlying long- and short-term fMRI adaptations. In addition, the bandwidth of tuning curves in both LO and rFFA suggested by the short-term adaptation paradigm is much narrower than that by the long-term adaptation paradigm and less consistent with the psychophysical and single-unit studies. A very recent study with the short-term adaptation paradigm (Sawamura and others 2006) found neuronal adaptation at the single-cell level showed a greater degree of stimulus selectivity than neuronal responses, which means that the neuronal tuning bandwidth estimated from short-term adaptation is narrower than when measured with traditional paradigms. Another fMRI short-term adaptation study (Gilaie-Dotan and Malach forthcoming) about face identity encoding also suggested high face sensitivity—and a very narrow tuning curve—in fusiform face area (FFA) (but psychophysical adaptation experiments [Leopold and others 2001; Webster and others 2004] suggested a broad tuning curve. See also Loffler and others 2005). These studies are consistent with our short-term adaptation results.

The most dramatic difference between these 2 paradigms is in the rSTS. The failure to find fMRI adaptation effect with short-term adaptation in rSTS underscores the need for caution when interpreting null effects in adaptation experiments. Though null effects are always difficult to interpret, they are frequently used in adaptation experiments to make claims about invariance. Indeed, it may reflect nothing but a negative result that could be better detected with longer adaptation durations. Overall, it is clear that potentially different mechanisms may support long- and short-term adaptations and that considering these different mechanisms have important implications for the inferences one can make from adaptation data. For example, as was discussed earlier, long-term adaptation may be more suitable for revealing feedforward neuronal sensitivity changes following long exposure to a certain stimulus, whereas short-term adaptation may reveal more of the feedback influences from higher level processing. This idea is supported by the observation that the effect of long-term adaptation could be seen in early components of event-related potentials (ERPs) (Heinrich and others 2005), suggesting that it affects feedforward sensitivity (Kohn and Movshon 2003). On the other hand, ERP and magneto-encephalographic studies suggest that the effect of short-term adaptation (or following brief exposure) usually shows up later in the dynamic response components (Dale and others 2000; Doniger and others 2001; Schendan and Kutas 2003; Henson and others 2004; Gruber and Muller 2005).

It should be noted that, in a typical short-term adaptation experiment, a large set of different stimuli are usually used and the adapting stimulus is changed on each trial. In the current study, for each adaptation scan, only one adapting stimulus and 2 other test stimuli were used, which may have led to cross-trial adaptation and resulted in less differential activation for the different test stimuli. The purpose of using limited stimuli was to match the experimental conditions between the long-term and short-term adaptations (following the psychophysical tradition, the same adapter needs to be used in both pre-adaptation and topping-up adaptation). However, our data suggest cross-trial adaptation is unlikely to have occurred. First, the BOLD signals released in the short-term adaptation experiment are comparable with those in the other studies (Kourtzi and Kanwisher 2000, 2001; Epstein and others 2003). Second, the BOLD signals released by presenting the Test30 are comparable between the long and short adaptation paradigms in FFA and are even larger with short-term adaptation in LO (see Figs 4 and 6). Again, we emphasize the potential mechanism difference but not the general adaptation strength difference.

Another difference between the long-term and short-term adaptation experiments is that the Fixation trials were different. It might be expected that the event-related BOLD signals in the long-term adaptation should be weaker than that in the short-term adaptation because the baseline in the former one is a floating face. However, after tens of seconds' adaptation, the floating face would not be a good stimulus for the face-selective areas, and functionally it is almost equivalent to a blank screen. Also, the onset of the following topping-up adapter might also have added some signal into the time courses in Figures 4 and 5. For adaptation experiments, the BOLD signals released from adaptation are more important for inferring neuronal selectivity.

In conclusion, the long-term fMRI adaptation paradigm has revealed a distributed viewer-center face representation in the human high-level visual cortex. Our study also revealed important differences between the short- and long-term adaptation

techniques. More long-term fMRI adaptation experiments are needed to understand adaptation mechanism(s) and face/object representations at a finer level.

Notes

Address correspondence to Fang Fang, Department of Psychology, University of Minnesota, 75 East River Parkway, Minneapolis, MN 55455, USA. Email: fang0057@umn.edu.

References

Andrews TJ, Ewbank MP. 2004. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. Neuroimage 23:905–913.

Anstis SM, Moulden BP. 1970. Aftereffect of seen movement: evidence for peripheral and central components. Q J Exp Psychol 22:222–229.

Avidan G, Hasson U, Hendler T, Zohary E, Malach R. 2002. Analysis of the neuronal selectivity underlying low fMRI signals. Curr Biol 12:964–972.

Biederman I. 1987. Recognition-by-components: a theory of human image understanding. Psychol Rev 85:249–277.

Blakemore C, Campbell FW. 1969. On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal image. J Physiol 203:237–260.

Boynton G, Finney E. 2003. Orientation-specific adaptation in human visual cortex. J Neurosci 23:8781–8787.

Buckner RL, Goodman J, Burock M, Rotte M, Koutstaal W, Schacter D, Rosen B, Dale AM. 1998. Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. Neuron 20:285–296.

Buckner RL, Koutstaal W. 1998. Functional neuroimaging studies of encoding, priming and explicit memory retrieval. Proc Natl Acad Sci USA 95:891–898.

Buracas GT, Boynton GM. 2002. Efficient design of event-related fMRI experiments using M-sequences. Neuroimage 15:801–813.

Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E. 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26:55–67.

Desimone R. 1996. Neural mechanisms for visual memory and their role in attention. Proc Natl Acad Sci USA 93:13494–13499.

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

De Souza WC, Eifuku S, Tamura R, Nishijo H, Ono T. 2005. Differential characteristics of face neuron responses within the anterior superior temporal sulcus of macaques. J Neurophysiol 94:1252–1266.

Doniger GM, Foxe JJ, Schroeder CE, Murray MM, Higgins BA, Javitt DC. 2001. Visual perceptual learning in human object recognition areas: a repetition priming study using high-density electrical mapping. Neuroimage 13:305–313.

Engel SA, Glover GH, Wandell BA. 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex 7:181–192.

Epstein R, Graham KS, Downing PE. 2003. Viewpoint-specific scene representations in human parahippocampal cortex. Neuron 37:865–876.

Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. Nature 392:598–601.

Fang F, He S. 2005. Viewer-centered object representation in the human visual system revealed by viewpoint aftereffects. Neuron 45:793–800.

Fang F, Murray SO, Kersten DJ, He S. 2005. Orientation-tuned fMRI adaptation in human visual cortex. J Neurophysiol 94:4188–4195.

Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. 2000. The fusiform "face area" is part of a network that processes faces at the individual level. J Cogn Neurosci 12:495-504.

Gilaie-Dotan S, Malach R. Sub-exemplar shape tuning in human face-related areas. Cereb Cortex. Forthcoming.

Grill-Spector K, Henson R, Martin A. 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci 10:14–23.

Grill-Spector K, Knouf N, Kanwisher N. 2004. The fusiform face area subserves faceperception, not generic within-category identification. Nat Neurosci 7:555-562.

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R. 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron 24:187-203.

Grossman ED, Blake R. 2002. Brain areas active during visual perception of biological motion. Neuron 35:1167–1175.

Gruber T, Muller MM. 2005. Oscillatory brain activity dissociates between associative stimulus content in a repetition priming task in the human EEG. Cereb Cortex 15:109-116.

Hasselmo ME, Rolls ET, Baylis GC, Nalwa V. 1989. Object centered encoding by face selective neurons in the cortex in the superior temporal sulcus of the monkey. Exp Brain Res 75:417–429.

Haxby JV, Hoffman EA, Gobbini MI. 2000. The distributed human neural system for face perception. Trends Cogn Sci 4:223-233.

Heinrich SP, Renkl AEH, Bach M. 2005. Pattern specificity of human visual motion processing. Vision Res 45:2137–2143.

Henson RN, Rylands A, Ross E, Vuilleumeir P, Rugg MD. 2004. The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. Neuroimage 21:1674–1689.

Kanwisher N, Yovel G. 2006. The fusiform face area: a critical region specialized for the perception of faces. Proc R Soc Lond B Biol Sci. Forthcoming.

Kohler W, Wallach H. 1944. Figural aftereffects: an investigation of visual processes. Proc Am Philos Soc 88:269-357.

Kohn A, Movshon JA. 2003. Neuronal adaptation to visual motion in area MT of the macaque. Neuron 39:681–691.

Kourtzi Z, Grill-Spector K. 2005. fMRI adaptation: a tool for studying visual representations in the primate brain. In: Rhodes G, Clifford CWG, editors. Fitting the mind to the world: adaptation and after-effects in high-level vision. Oxford: Oxford University Press. p 173-188.

Kourtzi Z, Kanwisher N. 2000. Cortical regions involved in perceiving object shape. J Neurosci 20:3310-3318.

Kourtzi Z, Kanwisher N. 2001. Representation of perceived object shape by the human lateral occipital complex. Science 293:1506–1509.

Krekelberg B, Boynton GM, van Wezel RJA. 2006. Adaptation: from single cells to BOLD signals. Trends Neurosci 29:250-256.

Larsson J, Landy MS, Heeger DJ. 2006. Orientation-selective adaptation to first- and second-order patterns in human visual cortex. J Neurophysiol 95:862-881.

Leopold D, O'Toole A, Vetter T, Blanz V. 2001. Prototype-referenced shape encoding revealed by high-level aftereffects. Nat Neurosci 4:89-94.

Leopold D, Rhodes G, Muller K, Jeffery L. 2005. The dynamics of visual adaptation to faces. Proc R Soc Lond B Biol Sci 272:897-904.

Loffler G, Yourganov G, Wilkinson F, Wilson HR. 2005. fMRI evidence for the neural representation of faces. Nat Neurosci 8:1386-1390.

Murray SO, Olman CA, Kersten DJ. 2006. Spatially specific fMRI repetition effects in human visual cortex. J Neurophysiol 95:2439-2445.

Murray SO, Wojciulik E. 2004. Attention increases neural selectivity in the human lateral occipital complex. Nat Neurosci 7:70-74.

Perrett DI, Harries MH, Bevan R, Thomas S, Benson PJ, Mistlin AJ, Chitty AJ, Hietanen JK, Ortega JE. 1989. Frameworks of analysis for the neural representation of animate objects and actions. J Exp Biol 146:87-113.

Perrett DI, Hietanen JK, Oram MW, Benson PJ, Rolls ET. 1992. Organization and functions of cells responsive to faces in the temporal cortex. Philos Trans R Soc Lond B Biol Sci 335:23-30.

Perrett DI, Oram MW, Harries MH, Bevan R, Hietanen JK, Benson PJ, Thomas S. 1991. Viewer-centered and object-centered coding of heads in the macaque temporal cortex. Exp Brain Res 86:159-173.

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

Perrett DI, Smith PAJ, Potter DD, Mistlin AJ, Head AS, Milner AD, Jeeves MA. 1984. Neurons responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. Hum Neurobiol 3:197-208.

Poggio T, Edelman S. 1990. A network that learns to recognize three-dimensional objects. Nature 343:263-266.

Pourtois G, Schwartz S, Seghier ML, Lazeyras F, Vuilleumier P. 2005. Portraits or people? Distinct representations of face identity in the human visual cortex. J Cogn Neurosci 17:1043-1057.

Quian Quiroga R, Reddy L, Kreiman G, Koch C, Fried I. 2005. Invariant visual representation by single neurons in the human brain. Nature 435:1102-1107.

Rhodes G, Jeffery L, Watson TL, Clifford CWG, Nakayama K. 2003. Fitting the mind to the world: face adaptation and attractiveness aftereffects. Psychol Sci 14:558-566.

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.

Schendan HE, Kutas M. 2003. Time course of processes and representations supporting visual object identification and memory. J Cogn Neurosci 15:111-135.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RBH. 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268:889-893.

Smith AM, Lewis BK, Ruttimann UE, Ye FQ, Sinnwell TM, Yang Y, Duyn JH, Frank JA. 1999. Investigation of low frequency drift in fMRI signal. Neuroimage 9:526-533.

Suzuki S, Grabowecky M. 2002. Evidence for perceptual "trapping" and adaptation in multistable binocular rivalry. Neuron 36:143-157.

Talairach J, Tournoux P. 1988. Co-planar stereotaxic atlas of the human brain. New York: Thieme Medical Publishers.

Ullman S. 1989. Aligning pictorial descriptions: an approach to object recognition. Cognition 32:193-254.

van Turennout M, Bielamowicz L, Martin A. 2003. Modulation of neural activity during object naming: effects of time and practice. Cereb Cortex 13:381-391.

Wagner AD, Gabrieli JD, Verfaellie M. 1997. Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. J Exp Psychol Learn 23:305-323.

Watson TL, Clifford CWG. 2003. Pulling faces: an investigation of the face-distortion aftereffect. Perception 32:1109-1116.

Webster MA, Kaping D, Mizokami Y, Duhamel P. 2004. Adaptation to natural facial categories. Nature 428:557-561.

Webster MA, Maclin LH. 1999. Figural aftereffects in the perception of faces. Psychon Bull Rev 6:647-653.

Wiggs CL, Martin A. 1998. Properties and mechanisms of perceptual priming. Curr Opin Neurobiol 8:227-233.

Zhao L, Chubb C. 2001. The size-tuning of the face distortion aftereffect. Vision Res 41:2979-2994.