Vision Research 49 (2009) 2800–2807

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Adaptation in the fusiform face area (FFA): Image or person?

Xiaokun Xu [a,*], Xiaomin Yue [b], Mark D. Lescroart [c], Irving Biederman [a,c], Jiye G. Kim [a]

[a] Department of Psychology, University of Southern California, Los Angeles, CA 90089-2520, USA
[b] Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02129, USA
[c] Neuroscience Program, University of Southern California, Los Angeles, CA 90089-2520, USA

ARTICLE INFO

Article history:
Received 18 March 2009
Received in revised form 19 August 2009

Keywords:
Face recognition
Fusiform face area
fMRI adaptation
Face representation
Gabor-jet scaling

ABSTRACT

Viewing a sequence of faces of two different people results in a greater Blood Oxygenation Level Dependent (BOLD) response in FFA compared to a sequence of identical faces. Changes in identity, however, necessarily involve changes in the image. Is the release from adaptation a result of a change in face identity, per se, or could it be an effect that would arise from any change in the image of a face? Subjects viewed a sequence of two faces that could be of the same or different person, and in the same or different orientation in depth. Critically, the physical similarity of view changes of the same person was scaled, by Gabor-jet differences, to be equivalent to that produced by an identity change. Both person and orientation changes produced equivalent releases from adaptation in FFA (relative to identical faces) suggesting that FFA is sensitive to the physical similarity of faces rather than to the individuals depicted in the images.

1. Introduction

The fusiform face area (FFA) is a region of the human ventral visual pathway that exhibits a greater BOLD response to faces than objects like tools, houses, and appliances (Kanwisher, McDermott, & Chun, 1997; Puce, Allison, Asgari, Gore, & McCarthy, 1996). However, whether the *identity* of a face, per se, is represented in this area remains unclear.

Some evidence is suggestive of a contribution of FFA to the perception of face identity. Experiments using adaptation paradigms have shown that sequential presentation of face images from different individuals produces higher activation in FFA, compared to the repetition of faces of the same person (Andrews & Ewbank, 2004; Eger, Schyns, & Kleinschmidt, 2004; Gauthier et al., 2000; Gilaie-Dotan & Malach, 2007; Loffler, Yourganov, Wilkinson, & Wilson, 2005; Rhodes & Jeffery, 2006; Winston, Henson, Fine-Goulden, & Dolan, 2004). This effect has been interpreted as a release from the adaptation produced by repetition of the identity of a face. Rotshtein, Henson, Treves, Driver, and Dolan (2005) further showed that the magnitude of the release from adaptation in FFA followed the perceived identity change, not the physical change along a series of morphed pictures between two celebrities' faces, for example Marilyn Monroe and Margaret Thatcher. In addition, activation in FFA has been shown to be higher for the successful identification of a particular face, comparing with the detection of a face vs. non-face categories (Grill-Spector, Knouf, & Kanwisher, 2004).

Consistent with the coding of individuation in FFA are the deficits produced by lesions of the occipito-temporal area, including FFA in prosopagnosics (Damasio, Tranel, & Damasio, 1990; Schiltz et al., 2006). These individuals are able to detect faces, but have difficulty in identifying them. They exhibit greater activation to faces compared to objects in FFA, but their activation in FFA is equivalent for conditions presenting identical and distinct faces, in contrast with the larger activation to distinct faces than identical ones in control subjects (Dricot, Sorger, Schiltz, Goebel, & Rossion, 2008; Schiltz et al., 2006).

However, the inference that FFA is representing identity is complicated by the finding that different poses of the same person also produce a release from adaptation (Andrews & Ewbank, 2004; Fang, Murray, & He, 2007; Pourtois, Schwartz, Seghier, Lazeyras, & Vuilleumier, 2005). FFA, defined by a contrast of faces minus objects, has also been shown to be sensitive to a number of variations of face stimuli, in addition to identity. In a block-design fMRI-adaptation (fMRI-a) experiment, epochs with either translation or rotation in depth of the same face, or a change of identity all produced greater BOLD responses than those with faces from the same person (Grill-Spector et al., 1999). Eger et al. (2004) showed that a change in face identity or spatial frequency band of a filtered face also produced equivalent greater activation in FFA than identical faces. Even more striking with respect to FFA's sensitivity to image variables is that complementary images in the Fourier domain of the same person—where members of a complementary pair would each contain every other combination of eight orientations and

* Corresponding author. Tel.: +1 213 740 6102; fax: +1 213 740 5687.
E-mail address: xiaokunx@usc.edu (X. Xu).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.08.021

eight scales—produced as large a release from adaptation as images of different people (Yue, Tjan, & Biederman, 2006). However, the physical change produced by the complementation (i.e., different frequency-orientation kernels), was greater than that produced by a change of person with the same frequency-orientation kernels (Yue et al., 2006), as assessed by the Gabor-jet measure (Lades et al., 1993). This was also true of the study of Grill-Spector et al. (1999) in that the image change produced by rotation in depth was greater than that produced by change in person in pixel-energy measurement.

To assess the effects of changes in person and physical image on adaptation in FFA, we designed a fast event-related fMRI adaptation experiment in which subjects viewed sequences of two faces that could vary in identity, orientation in depth, both, or neither. A critical feature of the design, vis-à-vis prior studies, was that the magnitude of the image change produced by a change in person was equivalent to the magnitude of the image change produced by a change in orientation (see Fig. 1) on a trial-by-trial basis, as assessed by the Gabor-jet model (Biederman & Kalocsai, 1997; Lades et al., 1993), a model of V1 cell filtering (see Section 2). In addition to having high merit as a computer model of face recognition (having won a US competition, Okada et al., 1998), the model captures many of the phenomena of face recognition and it predicts psychophysical similarity of face discriminability almost perfectly (Yue, Subramaniam, & Biederman, 2007). Specifically, on a match-to-sample task in which subjects had to determine which one of two faces (a matching face and a distracter) was identical to a sample face, the correlation between the physical similarity of the distracter (as scaled by Gabor jets) to the matching face and error rates was in the high .90 s.

Given the prior studies, a change in viewpoint would be expected to result in a release of adaptation in FFA (e.g., Fang et al., 2007). The critical issue under test in the present investigation is whether there would be a greater release from adaptation if there was a change of person when the magnitude of such image changes were matched to viewpoint changes according to the Gabor-jet model, on a trial-by-trial basis.

2. Methods

2.1. Subjects

Seventeen subjects (six females), mean age 26 ± 1.2 years, participated in the experiment. All subjects reported normal or corrected-to-normal vision and had no known neurological or visual disorders. They gave written informed consent prior to the experiment, in which all procedures and protocols were approved by the Institutional Review Board of the University of Southern California.

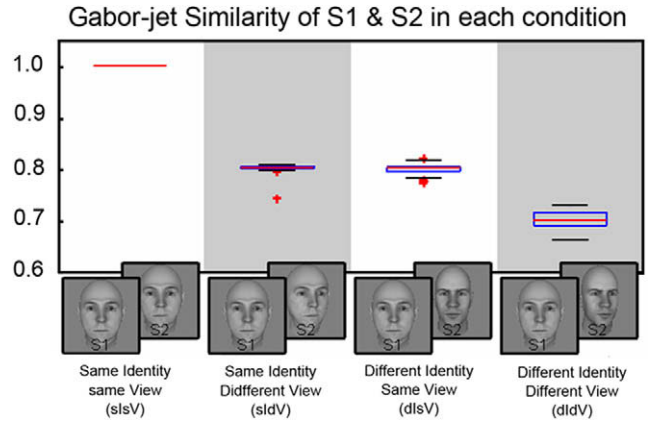

Fig. 1. Sample stimuli for the four conditions (here illustrated with the same S1) and a boxplot of the Gabor-jet similarity values for each condition. The plus signs represent outliers beyond the 1.5 interquartile range. In sIdV, both outliers are below the mean.

2.2. Magnetic resonance imaging setup

Subjects were scanned in a 3.0-T Siemens MAGNETOM Trio Scanner equipped with a 12-channel head coil at the Dana and David Dornsife Cognitive Neuroscience Imaging Center at the University of Southern California.

2.3. Stimuli

All stimuli were generated by the FaceGen Modeller 3.2 (http://www.facegen.com). The face models were chosen to be middle aged, Caucasian males, without hair on a gray background (Fig. 1). Identity was varied by modification of both local features (the shape of the eyes, nose, mouth, jaw and cheekbone) and their spatial configuration, for example the distance between the eyes, nose and mouth. The rotated orientation was approximately 20° to the right from the frontal orientation to allow equivalent Gabor-jet similarity scaling of orientation and identity changes. The rotation angle was estimated by exporting the model into 3dsMax (Autodesk, http://www.usa.autodesk.com). Finally, all the stimuli were sized to large (6° × 6°) and small (3° × 3°) versions (for the behavioral task).

The combination of identity (same vs. different) and viewpoint (same vs. different) between two faces in each pair resulted in four conditions: Same Identity Same View (sIsV), Same Identity Different View (sIdV), Different Identity Same View (dIsV), and Different Identity Different View (dIdV). The mean and standard error of the Gabor-jet similarity values (explained below) between the pair of faces in each trial was 1.00 ± 0 for sIsV, 0.80 ± 0.001 for both sIdV and dIsV, and 0.70 ± 0.001 for dIdV condition, as shown in Fig. 1. Paired *t*-tests showed that the similarity for dIdV was lower than both sIdV: $t(49) = 39.7$, $p < 0.001$ and dIsV conditions: $t(49) = 34.9$, $p < 0.001$, but the similarity for the sIdV and dIsV conditions was not different from each other: $t(49) = 1.4$, $p = 0.16$.

Subjects were instructed to judge whether the images on a given trial were the same or different in size. By having the task independent of the variables of interest (pose and person), we reduced the likelihood that attentional strategies tuned to individuation or pose would modulate the BOLD response.

2.4. Gabor-jet similarity scaling

The Gabor-jet similarity value (Fiser, Biederman, & Cooper, 1996; Lades et al., 1993) for each pair of stimuli was computed from a 10 × 10 grid centered on each picture with each node of the grid corresponding to the center of the receptive field of one jet. Each jet was composed of 40 Gabor filters (or kernels) at eight equally spaced orientations (i.e., 45° differences in angle) × 5 scales, each centered on their jet's grid point. The coefficients of the kernels (the magnitude corresponding to an activation value for a complex cell) within each jet were then concatenated to a 4000-element (100 jets × 40 kernels) vector G: [$g1$, $g2, \ldots, g4000$]. For any pair of pictures with corresponding jet coefficient vectors G and F, the similarity of the pairs was defined as:

$$Sim(G,F) = \frac{\sum_{i=1}^{4000} g_i f_i}{\sqrt{\sum_{i=1}^{4000} g_i^2} * \sqrt{\sum_{i=1}^{4000} f_i^2}}, i = 1, 2, \ldots, 4000$$

This is a correlation between the vectors (corresponding to the cosine of their angular difference) and yields a similarity value between 0 and 1.00.

2.5. Scanning parameters

For functional scanning, BOLD contrast was obtained with a gradient-echo echo planar imaging (EPI) sequence. The parameters were: TR = 2 s; TE = 30 ms; flip angle = 60°; field of view = 224 × 224 mm, matrix size = 64 × 64, in-plane resolution = 3.5 mm, slice thickness = 3 mm, between-slice gap = 0. The scanning volume consisted of continuous 32 slices covering most of the lower cortex, including the temporal poles. For anatomical scanning, a whole brain three-dimensional T1-weighted structural scan was done with MPRAGE sequences, with parameters as follows: TI = 1100, TR = 2.07 s, TE = 4.1 ms, Flip angle = 12°, 192 sagittal slices, matrix size = 256 × 256, voxel resolution = 1 × 1 × 1 mm.

2.6. Localizer runs

To define face- and object-selective Regions of Interest (ROIs), subjects passively viewed blocks of grayscale pictures of faces, objects, and scrambled textures of the faces and objects. The faces were randomly selected from the Stirling University face database (http://www.pics.psych.stir.ac.uk) with equal numbers of male and female faces, all in frontal views with neutral expressions. The objects were cartoons of common household objects. Texture was created by scrambling 8 by 8 pixel patches of the intact images so that no discernible features were apparent. The localizer run consisted of 15 12 s-blocks. Each block included 24 different pictures from one of the three categories, presented for 500 ms sequentially, and each category block was repeated five times, in randomized order across the session. Stimuli were presented in the center of screen and subtended a visual angle of 7° × 7°. During the scanning, subjects were asked to maintain center fixation when viewing the stimuli.

2.7. Event-related adaptation design

Four fast event-related fMRI-adaptation scans were used to test the sensitivity of FFA to changes in identity and/or viewpoint. In a given scan, a subject viewed a sequential pair (S1, S2) of faces on each trial, and responded on an MRI-compatible button box to indicate whether or not the two faces were the same or different in size (either small or large), a judgment that was independent of the identity and pose of the faces. The duration for each trial was 2 s. S1 was presented for 300 ms, followed by a 400 ms blank screen with fixation, and then S2 was presented for 300 ms followed by a 1 s blank during which the subject responded. The timing parameters were the same as those used by Kourtzi and Kanwisher (2000) and Winston et al. (2004). No feedback was provided in the actual scanning sessions. Each run had a total duration of 8 min 44 s, consisting of 252 trials. There was an initial 10 s fixation period with a black dot centered on a gray background screen to compensate for the initial magnetic field inhomogeneity and a final 10 s fixation period to accommodate the lag of the hemodynamic response at the end of the scan session. Before going into the scanner, subjects were given 100 practice trials, using a different set of stimuli. During the practice trials, feedback was provided for incorrect responses, in which a red dot appeared when the subject's response was in error, or a yellow dot appeared when the subject's response missed the response interval.

Each run was composed of 50 trials per Identity–View condition plus 52 catch trials with a blank screen throughout the run. For each condition, the four possible S1-S2-size-configurations (small–small, small–large, large–small, large–large) were equally distributed within all trial types (sIsV, sIdV, dIsV, dIdV) and collapsed in the analysis. Ordering of the conditions was arranged such that the history of the two preceding trials for each trial was composed of equal numbers of all the conditions including the 4 experiment condition as well as fixation trials. Therefore, the inter-stimulus-interval was jittered between 2 s to 6 s.

3. Data analysis

The imaging data were analyzed with Brainvoyager QX (Brain Innovation BV, Masstricht, Netherlands). All data from a scan were preprocessed with 3D motion-correction, slice timing correction, linear trend removal and temporal smoothing with a high pass filter set to three cycles over the run's length. Both an ROI and a multi-voxel analysis were conducted. A 4 mm Gaussian kernel was also used in the spatial smoothing of the functional images prepared for ROI analysis in face-selective areas, whereas the functional images entering multi-voxel analysis were not smoothed to maintain the intrinsic fine-grained pattern of local differences which is critical to the multivariate approach (Kriegeskorte & Bandettini, 2007). Each subject's preprocessed image was then coregistered with their same-session, high-resolution anatomical scan. Then each subject's anatomical scan was transformed into Talairach coordinates. Finally, using the above transformation parameters, the functional image was transformed to Talairach coordinates as well. All statistical analysis was performed on the transformed data.

Face-selective regions were defined as regions with greater activation to intact faces than to objects. The threshold of voxel activation was set at Bonferroni corrected $p < 0.05$ for each subject, for the contrast of faces minus objects. The threshold of activation extension was set as 30 continuous voxels. We also defined a non-face-selective area, the Lateral Occipital Complex (LOC), based on a contrast of objects minus scrambled objects (i.e., texture), with the same statistical threshold. In addition, an early visual area, based on a contrast of texture minus fixation, was anatomically localized around the calcarine sulci.

For the event-related experimental scans, a deconvolution analysis was performed on all voxels within each subject's localizer-defined ROI to estimate the time course of the BOLD response, for each trial type. Deconvolution was computed by having ten 2 s shifted versions of the indicator function for each trial type and response type (correct or incorrect) as the regressor in a fixed-effect general linear model. The percent signal change was computed as the beta values for each regressor, divided by the mean activation values (value of beta zero in the general linear model) of the whole ROI.

A 2 × 2 × 2 (Identity (same/different) × Viewpoint (same/different) × Size (same/different)) repeated measures analysis of variance (ANOVA) was performed on the peak BOLD response on correct responses. Inclusion of error trials did not change the pattern of the results. The average of the peak values of the hemodynamic curves, namely the mean percent signal change of the 3rd and 4th TR point (5–8 s from the onset of each trial) were computed for each condition (sIsV, sIdV, dIsV and dIdV with same size or different size, respectively). Reaction times (RTs) and percent correct were analyzed in a similar 2 × 2 × 2 ANOVA. Although the design fully crossed identity and viewpoint, the critical comparison in our investigation is not whether there would be a release from adaptation from a change of identity or viewpoint per se, but whether there would be a greater release from a change in person than a change in viewpoint, given that the low-level image property was matched in terms of Gabor-jet similarity. This crucial comparison was assessed by a post hoc paired *t*-test.

4. Results

4.1. Behavioral results for the same-different size judgment task

A change in Identity or View both interfered with the same-size response, producing longer RTs and higher error rates than the identical condition (Table 1).

To test the reliability of whether the different response type (same or different size) interacted with Identity and viewpoint

Table 1
Behavioral results (RTs and % correct) as a function of Identity, Viewpoint and Size (n = 17).

	Same-size S1 and S2				Different-size S1 and S2			
	sIsV	sIdV	dIsV	dIdV	sIsV	sIdV	dIsV	dIdV
RTs (ms)	580	595	608	613	612	601	611	610
S.E.M	22.6	22.8	22.1	20.4	20.8	21.3	20.4	20.4
% Correct	81.4	80.1	80.0	75.8	78.2	80.0	79.5	79.5
S.E.M	1.33	1.52	1.71	1.72	1.64	1.58	1.54	1.21

change, a 2 Size × 2 Identity × 2 View ANOVA was performed on the RTs and percent correct scores. RTs showed a significant main effect for Identity $F(1, 16) = 20$, $p < 0.001$ and significant interaction between Size and Identity, $F(1, 16) = 13.2$, $p < 0.01$, and Size and Viewpoint: $F(1, 16) = 6.1$, $p < 0.03$. For accuracy, none of the main effects were significant (all ps > 0.2), but interactions between Size and Viewpoint, $F(1, 16) = 8.0$, $p < 0.01$, and Size and Identity, $F(1, 16) = 3.1$, $p = 0.09$ were observed. We therefore broke down the trials into same-size response and different-size response groups, and performed the two-way (2 Identity × 2 View) ANOVA on each group.

For same size trials, the main effect of both Identity $F(1, 16) = 31.4$, $p < 0.01$ and Viewpoint $F(1, 16) = 8.6$, $p < 0.01$ were significant on RTs and accuracy, $F(1, 16) = 3.6$ $p < 0.05$, for identity, and $F(1, 16) = 9.8$, $p < 0.01$, for Viewpoint. For trials with different size faces, the main effects of trial type on RTs and accuracy were not significant, both $Fs(1, 16) < 1$.

In sum, even though the task did not require the subject to judge view or identity, there were, nonetheless, costs of changes in Identity or View on performance when the faces were the same sizes. That is, the detection of a difference in identity or View caused interference in judging that the sizes were the same. This inability to ignore the information in a face is consistent with the findings of automatic face processing without focal attention (Reddy, Reddy, & Koch, 2006) and automatic ultra fast saccades to face stimuli (Honey, Kirchner, & VanRullen, 2008).

4.2. Neuroimaging results

4.2.1. Region of interest localization

Fig. 2 shows the activation pattern for the contrast of face minus object from one subject superimposed on the anatomical image in Talairach coordinates. The face-selective ROI was consistently found in the right hemisphere for all subjects, whereas bilateral activation was only observed in five subjects. Talairach coordinates for the most consistent activation in the right hemisphere were X: 35 ± 0.6; Y: −50 ± 1.9; Z: −14 ± 0.8, with an activation size of 460 ± 64 mm^3, corresponding to the location of FFA in previous research (Grill-Spector et al., 2004; Fang et al., 2007). Activation in the right superior temporal sulcus (rSTS) was also found in 10 of the 17 subjects (as shown in Fig. 2) with the same contrast. The average Talairach coordinates for rSTS were X: 44 ± 1.1; Y: −44 ± 1.9; Z: 12 ± 0.7, with an activation size of 576 ± 101 mm^3. Other face-selective areas noted in previous studies (Fang et al., 2007; Yue et al., 2006), such as bilateral OFA, and left FFA, were found in only a few subjects.

The object-selective area — bilateral lateral occipital complex (LOC) — was localized by the contrasts of object minus scrambled texture in every subject. The average Talairach coordinates for the LOC were X: −39 ± 1.0; Y: −73 ± 2.5; Z: −4 ± 1.5, with an activation size of 5800 ± 680 mm^3 in the left hemisphere, and X: 33 ± 1.0; Y: −73 ± 2.5; Z: 3 ± 1.5, with an activation size of 5000 ± 1000 mm^3 in the right hemisphere.

As a reference, the activation pattern in early visual cortex was examined. This area was defined as a region anatomically located in the calcarine sulcus and functionally activated by texture (scrambled objects) vs. fixation trials at $p < 0.0001$ uncorrected threshold. The average Talairach coordinates of the Individually localized ROI for each subject were: X: −10 ± 1.5; Y: −86 ± 1.1; Z: −5 ± 1.4, with an activation size of 970 ± 107 mm^3 in the left hemisphere, and X: 7 ± 0.9; Y: −89 ± 1.1; Z: −4 ± 1.5, with an activation size of 895 ± 100 mm^3 in the right hemisphere. The anatomical loci are consistent with those in a review of early visual area localization by Hasnain, Fox and Woldorff (1998).

4.2.2. Region of interest analysis

A three-way ANOVA of percent BOLD signal change in response to Identity (same-different) × View (same-different) × Size (same-different), performed separately for each of the three ROIs (rFFA, LOC, and rSTS), revealed neither a main effect of Size, nor any interaction of Size with Identity and/or Viewpoint, all $Fs < 1$.

We therefore collapsed the size-variation across the Identity/Viewpoint conditions and ran a repeated measures 2 (Identity) × 2 (View) ANOVA for each ROI. For every subject, the change either in identity or viewpoint produced a greater BOLD response in right FFA, compared to the response to identical faces yielding significant main effects of both Identity $F(1, 16) = 7.6$, $p = 0.01$ and Viewpoint $F(1, 16) = 4.2$, $p = 0.05$, but no reliable interaction between the two factors $F(1, 16) = 1.4$, $p > 0.25$. A post hoc paired t-test, showed that the BOLD response in the sIsV condition was significantly smaller ($p < .05$) than each of the other three conditions. However, the release of adaptation for the three trial types did not differ from each other (all paired t-tests $t < 1$). In particular, a change of person did not produce a greater release from adaptation than a change in viewpoint. (see Fig. 3).

Only 10 of the 17 subjects showed a greater BOLD response to faces than objects in rSTS (Fig. 4). Unlike the pattern in FFA, the dIsV and sIdV conditions were equivalent to the sIsV condition but the dIdV condition had a larger BOLD release than the other three conditions. For these subjects there was no significant main

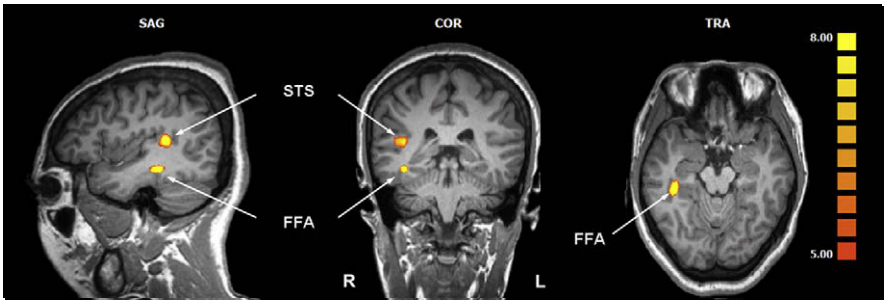

Fig. 2. The activation of face-selective ROIs in a typical subject's Talairach normalized brain, for the contrast of face minus objects (Right is left in the fMRI image. The threshold for the activation t-map is $p < 0.05$, Bonferroni corrected).

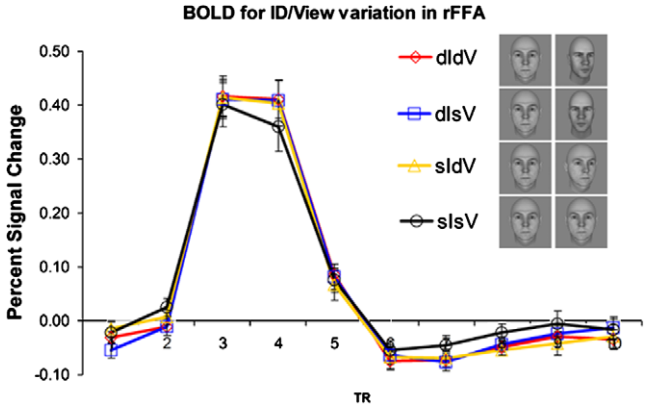

Fig. 3. Event-related BOLD response (percent signal change over fixation baseline) averaged over all subjects ($n = 17$) in right FFA.

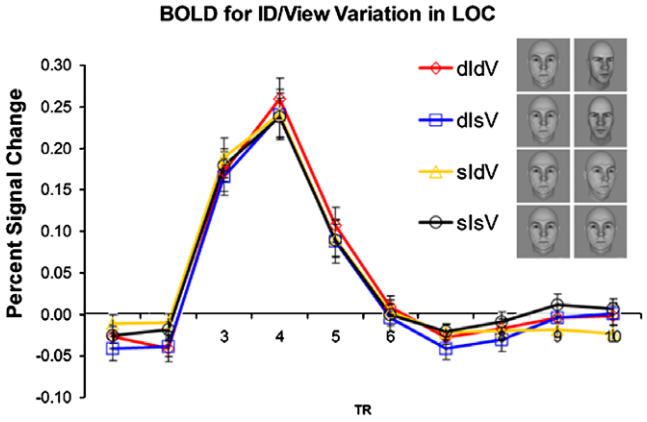

Fig. 5. Event-related BOLD response (percent signal change over fixation baseline) averaged over all subjects ($n = 17$) in bilateral LOC.

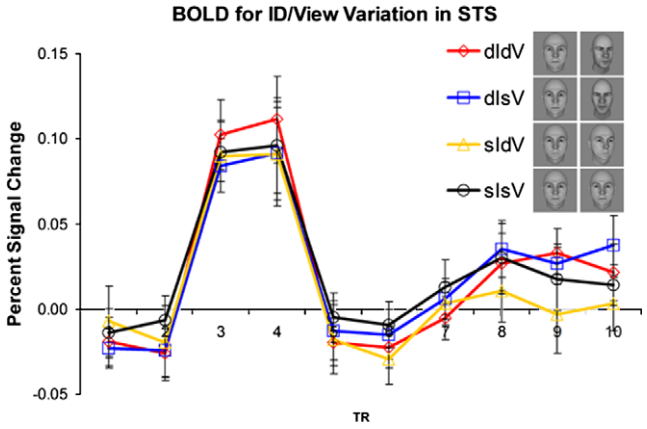

Fig. 4. Event-related BOLD response (percent signal change over fixation baseline) averaged over the 10 subjects (out of 17) who showed a greater BOLD response to the faces minus objects localizer in rSTS.

effect of either Identity or Viewpoint $Fs(1, 9) < 1$, and the interaction fell short of significance, $F(1, 9) = 1.6$, $p > 0.2$.

A 2×2 ANOVA in bilateral lateral occipital complex (LOC) revealed no significant effects of either Identity or Viewpoint: $F(1, 16) < 1$, $F(1, 16) = 1.9$, $p > 0.2$, respectively, or their interaction, $F(1, 16) < 1$ in both hemispheres (see Fig. 5).

In early visual cortex, the main effect of image similarity across the four conditions was significant, $F(3, 48) = 3.0$, $p < 0.05$. However, unlike the results for FFA, here there was a significantly higher percent signal change (4th TR point) in the dIdV condition (PSC = 0.110) than those in sIdV (PSC = 0.081), dIsV (PSC = 0.081) and sIsV (PSC = 0.085), all paired-tests $p < 0.05$. Again, the dIsV and sIdV conditions did not differ significantly from each other, $p > 0.5$. Overall, although there were some areas where larger Gabor scaled differences between stimuli did not result in a larger BOLD response, in no area was there ever a reversal of the ordering expected form Gabor similarity.

4.2.3. Multi-voxel analysis: Are different voxels selective for individuation and pose?

In an ROI analysis, the average time course across all the voxels within the ROI is used in the deconvolution general linear model. However, such averaging obscures the possibility that different voxels within the ROI—the FFA in the present case–are differentially selective for individuation and pose. Indeed, inhomogeneity within face-selective areas has been reported (Grill-Spector, Sayres, & Ress, 2006; Haxby et al., 2001). To address this hypothesis, we modeled each voxel (in functional images with a resolution of $3.5 \times 3.5 \times 3$ mm) within FFA separately, using the same general linear model that we used for the ROI analysis. We then measured the voxel-wise correlations of percent signal change between conditions. For all subjects, all conditions were very highly correlated with each other (all $rs > .90$) over the voxels. In an ANOVA of voxel-wise correlations, no significant differences were observed among the magnitude of correlations: $F(5, 80) = 1.5$, $p > .2$. Critically, there was no difference between the correlation of sIsV and dIdV and that of sIdV and dIsV. If there was independent coding of Person and View at the voxel level, a lower correlation should have been observed for sIdV and dIsV vectors. We should note, however, that this analysis had low power because the adaptation effects were small compared to the general variation in BOLD amplitude across the voxels, leading to overall high correlations. More importantly, the correlation analysis was limited in that both the distribution of the two types of (putative) neuron populations at the sub-voxel level and their equivalence in the magnitude of adaptation could profoundly affect the correlation pattern (Andresen, Vinberg, & Grill-Spector, 2009). We thus cannot, with confidence, reject the hypothesis that individual voxels are exclusively tuned to person and orientation.

4.2.4. Effect of absolute size

Although it was not a primary objective of this investigation, we were able to evaluate the effect of absolute face size in the different ROIs through the size manipulation for the subjects' task. Some caution in the results of this analysis is warranted, as we did not explicitly balance the various conditions with the size manipulation although we have no reason to think that there would be any systematic bias. As can be seen in Fig. 6 in early visual cortex (panel a), the larger the mean size of a pair of faces, the larger the BOLD response, $F(3, 48) = 53.63$, $p < 0.001$ over TRs 3 and 4). The ordering in that area was LL > LS = SL > SS (size of s1 and s2, Large or Small), both inequalities significant at $p < .001$. In later visual areas, although the overall effects of size was significant [$F(3, 38) = 6.33$ and 7.42, $p < .001$, in FFA and LOC, respectively, and $F(3, 27) = 4.15$, $p < .02$, in STS], with the largest faces consistently producing the largest BOLD response ($p < .05$ in FFA and LOC; ns in STS), the consistent positive association between face size and BOLD response evident in early cortical visual areas was not maintained.

Low experimental power could be a possible explanation for the lack of an effect of the conditions in bilateral LOC and rSTS. To rule out this alternative, we conducted a voxel-wise search in these ROIs. Even under a liberal threshold ($P < 0.001$ uncorrected) to allow any detection of difference among conditions, only a minimal number of voxels were detected (3 out 17 subjects had 4% voxel of their entire STS showing any difference among the conditions, and 4 out 17 subjects had less than 0.5% of the voxels of all of LOC showing a difference between dIdV and any of the other three conditions). Lowering the threshold of the localization did not change the pattern of results in both ROIs. Therefore, the absence of adaptation effect in these areas was not likely to be accounted for by low experimental power but, instead, a general insensitivity of LOC to (modest) variations in faces.

5.3. What role might FFA play in the face-processing system?

FFA reveals sensitivity to physical variations of faces—a phenomenon not characteristic of LOC—although the sensitivity to individuation is not any greater than that for other face variables, such as orientation in depth and frequency-orientation Fourier combinations. What role might FFA play in the face-processing system? At this point we can only speculate as to the functionality of FFA with respect to individuation. One possibility is that FFA is computing both face individuation and pose. If these were computed by separate subpopulations of neurons in FFA, we might have expected to see a greater release from adaptation when both person and pose were changed in the dIdV condition. However, such an additive effect was not observed in the ROI analysis. Nor did we observe any effects in the multi-voxel analysis that would be supportive of such separate subpopulations of neurons (although this analysis had low power).

An alternative possibility is that the adaptation of face-selective neurons in FFA is modulated by the simple physical similarity between faces: sIdV and dIsV had equivalent Gabor-jet similarity values and their release from adaptation was also equivalent in magnitude. This interpretation is compatible with results from both fMRI (Loffler et al., 2005; Rhodes & Jeffery, 2006) and single unit recoding in macaque IT (Leopold, Bondar, & Giese, 2006) in which the BOLD response in FFA and neuron firing rates in IT were proportionally modulated by the face's distinctiveness from the norm-face (mean face).

5.4. Effect of familiarity of face stimuli

The faces used in present study were unfamiliar to the subjects. The recognition or matching of unfamiliar faces shows greater costs when the images are changed compared to familiar faces, which can be identified even under low visual quality (Hancock, Bruce, & Burton, 2000). Could the release from adaptation in the sIdV condition be a result of the subject's failure to recognize that the two images were of the same person? To rule out this hypothesis, we asked the subject to perform an identification task after the scanning session (whether the same or different person was depicted in two stimuli in each trial) on a separate set of stimuli that were similar to those in the main experiment. The error rate on this task was less than 5%.

The present study indicates that the representation of unfamiliar faces in FFA is closely tied to the physical image and is, therefore, necessarily orientation sensitive. This is consistent with Ewbank and Andrews' (2008) finding that the priming effect in a behavioral identity-matching task and related fMRI adaptation study was also viewpoint-dependent for unfamiliar faces, but not for familiar ones. Similarly, the variation in faces of the same person either through sub-exemplar morphing (two faces perceived as the same person) (Gilaie-Dotan & Malach, 2007) or through changes of external features such as hairstyle (Davies-Thompson, Gouws, & Andrews, 2009), produced a release from adaptation relative to identical images of faces. In contrast, the representation of familiar faces was more invariant to image changes in FFA. Specifically, the release from adaptation in FFA by a change of person depicted in blocks of facial stimuli generalized across various viewpoints (Ewbank & Andrews, 2008). In another study using stimuli from the morphing continuum between two celebrities (Rotshtein et al., 2005), the release of adaptation in FFA paralleled the categorical perception of face identity, but not the within-category image changes.

5.5. A ceiling effect in fMRI adaptation?

The absence of an additive effect in the release of adaptation could be a simple consequence of the separate conditions of individuation and pose being at ceiling. Two face fMRI-a studies also found ceiling effects. In an event-related fMRI-a experiment, Fang et al. (2007) found that 60° and 90° rotations in depth of face stimuli did not produce any greater release in rFFA than a 30° rotation. Facial images perceived as belonging to the same individual (<35% in a morphing continuum between two people) were sufficient to produce full recovery in rFFA from adaptation (Gilaie-Dotan & Malach, 2007), equivalent to the release induced by images that were perceived as belonging to different (unfamiliar) people. This ceiling effect should not have been found in our ROI analysis if the identity and viewpoint of faces were coded independently in separate subpopulations of neurons at the voxel level. Instead, the data were more consistent with a model that assumes that the BOLD response in rFFA is modulated simply by the physical similarity between faces. The neural mechanism underlying adaptation could be the narrower turning and sparser representation in the distribution of neurons selective to face images (Grill-Spector, Henson, & Martin, 2006).

6. Conclusions

What might be the role of FFA in the face-processing system, given our result that sensitivity to identity changes in this area is not any greater than that for pose? It is possible that FFA is primarily serving as a face vs. non-face gate, passing on the image information relevant to individuation to a later area where individuation is made explicit. This information might be the spatial frequency and orientation content as suggested by Yue et al. (2006) or it could be the fragments suggested by Nestor, Vettel, and Tarr (2008). The area where individuation would actually be accomplished might be expected to be closer to associative cortex (Kriegeskorte, Formisano, Sorger, & Goebel, 2007), where units coding perceptual individuation could be linked to associative knowledge about the person, such as his or her profession, nationality, and name. The FFA gate might serve to protect these later face areas from non-face activity, a result that is highly compatible with Moeller, Freiwald, and Tsao's (2008) finding that face and non-face areas in the macaque were highly segregated. Perception or microstimulation of some face areas produced activation only in other face areas; none of the stimulation produced activation in non-face areas. Similarly, perception or microstimulation of non-face areas produced no activity in face areas. FFA may thus be an initial "protector" of later face networks. Given that so much of the image variation required for individuation of faces is subtle, it may be best to restrict the inputs to these later face areas so that the only inputs that affect connection weights of these face individuation networks are faces. This may be why prosopagnosics can show normal activation of FFA in that they know that a stimulus is a face (e.g.,

Schiltz et al., 2006). So, although perhaps necessary for individuation, FFA does not accomplish individuation.

Acknowledgments

We thank Kenneth Hayworth for helpful comments and Jiancheng Zhuang for his role in maintaining the MRI scanner. Supported by NSF 0420794, 0531177, and 0617699 to IB. The authors declare that they have no competing financial interests.

References

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage, 23*, 905–913.

Andresen, D. R., Vinberg, J., & Grill-Spector, K. (2009). The representation of object viewpoint in human visual cortex. *NeuroImage, 45*, 522–536.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society London: Biological Sciences, 352*, 1203–1219.

Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience, 13*, 89–109.

Davies-Thompson, J., Gouws, A., & Andrews, T. J. (2009). An image-dependent representation of familiar and unfamiliar faces in the human ventral stream. *Neuropsychologia, 47*, 1627–1635.

Dricot, L., Sorger, B., Schiltz, C., Goebel, R., & Rossion, B. (2008). The roles of "face" and "non-face" areas during individual face perception: Evidence by fMRI adaptation in a brain-damaged prosopagnosic patient. *NeuroImage, 40*, 318–332.

Ewbank, M. P., & Andrews, T. J. (2008). Differential sensitivity for viewpoint between familiar and unfamiliar faces in human visual cortex. *NeuroImage, 40*, 1857–1870.

Eger, E., Schyns, P. G., & Kleinschmidt, A. (2004). Scale invariant adaptation in fusiform face-responsive regions. *NeuroImage, 22*, 232–242.

Fang, F., Murray, S. O., & He, S. (2007). Duration-dependent FMRI adaptation and distributed viewer-centered face representation in human visual cortex. *Cerebral Cortex, 17*, 1402–1411.

Fiser, J., Biederman, I., & Cooper, E. E. (1996). To what extent can matching algorithms based on direct outputs of spatial filters account for human shape recognition? *Spatial Vision, 10*, 237–271.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience, 12*, 495–504.

Gilaie-Dotan, S., & Malach, R. (2007). Sub-exemplar shape tuning in human face-related areas. *Cerebral Cortex, 17*, 325–338.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*, 14–23.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience, 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.

Grill-Spector, K., Sayres, R., & Ress, D. (2006). High-resolution imaging reveals highly selective nonface clusters in the fusiform face area. *Nature Neuroscience, 9*, 1177–1185.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences, 4*, 330–337.

Hasnain, M., Fox, P., & Woldorff, M. (1998). Intersubject variability of functional areas in the human visual cortex. *Human Brain Mapping, 6*, 301–315.

Haxby, J., Gobbini, M., Furey, M., Ishai, A., Schouten, J., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science, 293*, 2425–2429.

Honey, C., Kirchner, H., & VanRullen, R. (2008). Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *Journal of Vision, 8*(9), 1–13.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience, 20*, 3310–3318.

Kriegeskorte, N., & Bandettini, P. (2007). Analyzing for information, not activation, to exploit high-resolution fMRI. *NeuroImage, 38*, 663–665.

Kriegeskorte, N., Formisano, E., Sorger, B., & Goebel, R. (2007). Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences, 104*, 20600–20605.

Lades, J. C. V., Buhmann, J., Lange, J., Malsburg, C., Wurtz, R., & Konen, W. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers: Institution of Electrical and Electronics Engineers, 42*, 300–311.

Leopold, D., Bondar, I., & Giese, M. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature, 442*, 572–575.

Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience, 8*, 1386–1390.

Moeller, S., Freiwald, W. A., & Tsao, D. Y. (2008). Patches with links: A unified system for processing faces in the macaque temporal lobe. *Science, 320*, 1355–1359.

Nestor, A., Vettel, J., & Tarr, M. (2008). Task-specific codes for face recognition: How they shape the neural representation of features for detection and individuation. *PLoS One, 12*, e3978.

Okada, K., Steffens, J., Maurer, T., Hong, H., Elagin, E., Neven, H., et al. (1998). The Bochum/USC face recognition system and how it fared in the FERET phase III test. In H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie, & T. Huang (Eds.), *Face recognition: From theory to applications (NATO ASI Series F)*. Berlin: Springer.

Pourtois, G., Schwartz, S., Seghier, M. L., Lazeyras, F., & Vuilleumier, P. (2005). Portraits or people? Distinct representations of face identity in the human visual cortex. *Journal of Cognitive Neuroscience, 17*, 1043–1057.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience, 16*, 5205–5215.

Reddy, L., Reddy, L., & Koch, C. (2006). Face identification in the near-absence of focal attention. *Vision Research, 46*, 2336–2343.

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research, 46*, 2977–2987.

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience, 8*, 107–113.

Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., et al. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. *Cerebral Cortex, 16*, 574–586.

Summerfield, C., Trittschuh, H. E., Monti, M. J., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience, 11*, 1004–1006.

Winston, J. S., Henson, R. N., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of Neurophysiology, 92*, 1830–1839.

Yue, X., Tjan, B. S., & Biederman, I. (2006). What makes faces special? *Vision Research, 46*, 3802–3811.

Yue, X., Subramaniam, S., & Biederman, I. (2007). Predicting the psychophysical discriminability of faces and other complex stimuli based on measures of physical image similarity. Presented at the Society for Neuroscience in San Diego, CA.