

# The Neural Code for Face Orientation in the Human Fusiform Face Area

Fernando M. Ramírez,<sup>1,2,3</sup> Radosław M. Cichy,<sup>1,2,3</sup>  Carsten Allefeld,<sup>1,3</sup> and John-Dylan Haynes<sup>1,2,3</sup>

<sup>1</sup>Bernstein Center for Computational Neuroscience, Charité-Universitätsmedizin Berlin, 10115 Berlin, Germany, <sup>2</sup>Berlin School of Mind and Brain, Humboldt Universität zu Berlin, 10099 Berlin, Germany, and <sup>3</sup>Berlin Center for Advanced Neuroimaging, Charité—Universitätsmedizin Berlin, 10115 Berlin, Germany

Humans recognize faces and objects with high speed and accuracy regardless of their orientation. Recent studies have proposed that orientation invariance in face recognition involves an intermediate representation where neural responses are similar for mirror-symmetric views. Here, we used fMRI, multivariate pattern analysis, and computational modeling to investigate the neural encoding of faces and vehicles at different rotational angles. Corroborating previous studies, we demonstrate a representation of face orientation in the fusiform face-selective area (FFA). We go beyond these studies by showing that this representation is category-selective and tolerant to retinal translation. Critically, by controlling for low-level confounds, we found the representation of orientation in FFA to be compatible with a linear angle code. Aspects of mirror-symmetric coding cannot be ruled out when FFA mean activity levels are considered as a dimension of coding. Finally, we used a parametric family of computational models, involving a biased sampling of view-tuned neuronal clusters, to compare different face angle encoding models. The best fitting model exhibited a predominance of neuronal clusters tuned to frontal views of faces. In sum, our findings suggest a category-selective and monotonic code of face orientation in the human FFA, in line with primate electrophysiology studies that observed mirror-symmetric tuning of neural responses at higher stages of the visual system, beyond the putative homolog of human FFA.

**Key words:** fMRI; face recognition; object recognition; invariance; orientation encoding; mirror symmetry

## Introduction

Object recognition is ecologically vital. Despite huge variation in stimulus properties, such as position, size, and orientation, our visual system meets the challenge within a fraction of a second (Thorpe et al., 1996). The computations involved and the supporting biological substrates remain a focus of intensive research (DiCarlo et al., 2012).

In recent decades, psychophysical evidence has accumulated indicating that human object recognition is view-dependent (Bülthoff and Edelman, 1992; Tarr et al., 1998; Fang and He, 2005), leading to computational theories based on 2D-view interpolation (Bülthoff et al., 1995; Ullman, 1998). Electrophysiological studies of object- and face-selective neurons in the ventral processing stream show that the vast majority of cells are unimodally tuned to a single preferred view (Perrett et al., 1991, 1998; Logothetis et al.,

1995). fMRI-guided electrophysiological studies recently revealed a system of face-selective patches along the macaque ventral stream (Tsao et al., 2003, 2006). Critically, Freiwald and Tsao (2010) found qualitatively distinct neural tuning functions to head orientation depending on the location of a face-patch within the ventral stream. The authors argue that intermediate regions of this system, exhibiting mirror-symmetrically tuned responses (e.g., similar responses to right and left profiles), might represent a key computational step before attaining full view-invariance. Interestingly, mirror-symmetrically responding patches were situated beyond the middle face patch (MFP), putative macaque homolog of human fusiform face-selective area (FFA) (Tsao et al., 2003).

Human neuroimaging studies have shown that face and object representations are viewpoint dependent, mostly revealing monotonic effects on brain responses as angular distance increases (Grill-Spector et al., 1999; Gauthier et al., 2002; Andrews and Ewbank, 2004; Fang et al., 2007b), reminiscent of the unimodal tuning functions observed in monkeys. However, two recent studies (Axelrod and Yovel, 2012; Kietzmann et al., 2012) combining fMRI and multivariate pattern analysis (MVPA) (Haynes and Rees, 2006; Norman et al., 2006) reported evidence of mirror-symmetric face encoding in several ventral visual areas, including FFA. However, these studies presented stimuli in a single location, such that low-level retinotopic effects could have influenced the observed effects. Also, although previous MVPA studies independently localized FFA, because they investigated

Received July 24, 2013; revised June 20, 2014; accepted July 11, 2014.

Author contributions: F.M.R., R.M.C., and J.-D.H. designed research; F.M.R. performed research; F.M.R. and C.A. contributed unpublished reagents/analytic tools; F.M.R. analyzed data; F.M.R., R.M.C., C.A., and J.-D.H. wrote the paper.

F.M.R. work was supported by a DAAD-CONICYT doctoral scholarship and by the Berlin School of Mind and Brain. R.M.C. was supported by a Feodor-Lynen scholarship of the Humboldt-foundation. We thank Jakob Heinzle, Robert Martin, Nikolaus Kriegeskorte, and Yi Chen for helpful comments and suggestions.

The authors declare no competing financial interests.

Correspondence should be addressed to Fernando M. Ramírez, Bernstein Center for Computational Neuroscience, Charité-Universitätsmedizin Berlin, Philippstr. 13, Haus 6, 10115 Berlin, Germany. E-mail: toporam@gmail.com.

DOI:10.1523/JNEUROSCI.3156-13.2014

Copyright © 2014 the authors 0270-6474/14/3412155-13\$15.00/0

orientation encoding only for faces, the category specificity of orientation information in this area remains unknown.

Here, we investigated the distributed responses associated with faces and vehicles under two transformations (rotation in-depth and retinal translation) in early visual cortex (EVC), lateral occipital cortex (LO) (Malach et al., 1995), and the right FFA (rFFA) (Kanwisher et al., 1997). We used a combination of fMRI, MVPA, and computational modeling to address the following questions: At what level of the visual hierarchy do object representations generalize across changes in position? Is orientation encoding in FFA category-selective? Does the observed representational structure derive from cells unimodally tuned to one preferred view or from mirror-symmetrical bimodal tuning curves? Finally, we substantiate our results and further investigate orientation encoding in FFA by means of a biologically informed computational model.

## Materials and Methods

### Participants

Nine healthy right-handed subjects (3 female, mean age 28 years, range 22–33 years) with normal or corrected-to-normal vision participated in the study. The study was approved by the ethics committee of the Max-Planck Institute for Human Cognitive and Brain Sciences (Leipzig) and conducted according to the Declaration of Helsinki. Informed consent was obtained from all participants. One subject was excluded from the study because rFFA could not be identified.

### Visual stimuli

We generated 3D meshes of faces/heads using FaceGen (Singular Inversions) and subsequently imported them into MATLAB (MathWorks). The two synthetically generated faces (male, female) were embedded in hairless head structures. We also used vehicle meshes corresponding to two Ford pickup trucks (model years 1928 and 1940). For rendering of stimuli, we used MATLAB in-house scripts, which provided tight control of lighting, position, size, and orientation. Stimuli were presented in the scanner using PsychToolbox (Brainard, 1997) and projected via an LCD projector (1024 × 768 pixels) onto a translucent screen at the rear of the scanner, and subjects saw them via a mirror mounted on the head coil (screen size: 25° × 18.75° of visual angle). Face and vehicle 3D meshes were rendered in grayscale in one of 10 combinations of position (3° above or below fixation) and orientation in-depth (−90°, −45°, 0°, 45°, 90°; front view = 0°) and subtended ~3.6° of visual angle (see Fig. 1). All objects had a uniform surface texture and were illuminated using the same 3 point lighting model. Illumination was kept identical for all objects under all conditions, except for a subtle luminance-level variation randomly introduced within each mini-block (for further details, see Experimental design).

### Localizers

To identify face- and object-selective cortical regions, each subject completed a combined FFA and LOC blocked localizer scan in the same session as the main experiment (Malach et al., 1995; Kanwisher et al., 1997). Subjects viewed blocks of images (duration: 16 s). Each block consisted of faces, common objects, or scrambled versions of these images. Blocks were presented in pseudo-random order and interleaved with periods of a uniform black background (duration: 16 s). Each block consisted of 20 images (700 ms per image, 100 ms gap). In each block, randomly four of the images were consecutively repeated. Subjects were asked to maintain fixation on a central dot and to perform a one-back task on image repetitions to sustain attention on the images. Subjects were asked to indicate repetitions via a button press.

### Main experiment

Subjects were instructed to maintain fixation on a red dot at the center of the screen (see Fig. 1A, right) and to perform a secondary task. While maintaining fixation, subjects were presented with mini-blocks (duration: 6 s) consisting of successive renderings (6 iterations, each displayed for 800 ms and followed by a 200 ms gap) of a single 3D mesh corresponding to one of two possible exemplars, belonging to two possible

categories (faces or vehicles). Within each mini-block, the 3D mesh was rendered in one of five angles of rotation in-depth (−90°, −45°, 0°, 45°, or 90°), and presented at one of two positions (3° above or below fixation). Thus, a total of 40 conditions were specified by combinations of the experimental factors (2 categories, 2 exemplars, 2 positions, 5 angles). Each of the 40 experimental conditions was presented twice per run. Subjects completed five runs of the main experiment. Each run was independently randomized. A 12 s fixation period was included before the presentation of the first mini-block and at the end of each run. Condition order and fixation periods (2–10 s) were scheduled using Optseq2 to optimize design orthogonality and efficiency. Total duration of each run was 745 s (298 volumes). To ensure attention to objects during the main experiment, participants were asked to indicate with a button-press at the end of every mini-block whether the number of subtly brighter presentations among the six displayed images was even (middle finger) or odd (index finger). Consistent with the instructions, subtle luminance differences were introduced among the six images presented within each mini-block, each image being displayed at one of two possible luminance levels. The number of brighter images within a mini-block was pseudo-randomly set to be either 3 or 4 of the 6 images, and the order of presentation of the images was also pseudo-randomized.

### Brain imaging

MRI data were acquired on a 3T Siemens Trio scanner with a 12-channel head coil. Structural images were acquired using a T1-weighted sequence (TR = 1900 ms, TE = 2.52 ms, flip angle = 9°, matrix size = 256 × 256, FOV = 256 mm, 192 sagittal slices of 1 mm thickness). Functional images were acquired with a gradient-echo EPI sequence (TR = 2500 ms, TE = 30 ms, flip angle = 70°, matrix = 128 × 96, FOV = 256 × 192 mm, 30 slices (2 mm thick, no gap, interleaved acquisition), resulting in a 2 mm isotropic voxel resolution. Slices were positioned along the slope of the temporal lobe and covered ventral visual cortex. A whole-brain EPI image (parameters as above but with 100 slices, TR = 8170 ms) was also collected each session. The functional localizer comprised 260 volumes and each of the runs of the main experiment 298 volumes.

### Data analysis

All functional data were analyzed using SPM2 (<http://www.fil.ion.ucl.ac.uk/spm>). EPI data were preprocessed and subsequently analyzed in single-subject space. That is, we did not perform spatial normalization of participants to a common reference space. For each subject, images were slice-time corrected, realigned to the first functional image, and resliced. Further analysis of fMRI data will be described in two parts. First, we will explain the analysis of the FFA and LOC localizer run and the selection of ROIs. Then, we will proceed to the analysis of the experimental data.

### Localization of ROIs

Functional data of the FFA and LOC localizer run were smoothed with a Gaussian kernel (FWHM = 5 mm). Only data used for localization of ROIs were smoothed, whereas experimental data remained unsmoothed for the purpose of multivariate analyses. We then modeled for each subject the cortical response of the localizer run with a GLM. The effects of intact objects, faces, and scrambled objects were modeled as three separate conditions. Regressors were convolved with an HRF.

**rFFA.** For each subject, we defined a face-sensitive region consisting of contiguous voxels in the right mid-fusiform gyrus (the FFA) that responded more strongly to faces than to objects at a significance level of  $p < 10^{-3}$  (uncorrected) at the voxel level (Kanwisher et al., 1997; Grill-Spector et al., 2004). In detail, we generated a  $t$  contrast “faces > objects” and included face-selective voxels within a sphere of 8 mm radius centered at the peak of the largest face-sensitive cluster ( $p < 0.05$ , FWE corrected; minimum cluster size = 20 voxels) observed in the fusiform gyrus (WFU PickAtlas fusiform mask inverse normalized to single subject space). These voxels were selected for the analysis of the experimental data.

**rOFA and rSTS.** We defined additional face-preferring areas using exactly the same criteria as rFFA, except for the necessary distinguishing anatomical criteria. Contiguous voxels labeled as OFA were located in inferior and/or middle occipital cortex, whereas rSTS was located in the STS and adjoining cortex of the superior or middle temporal cortex (WFU PickAtlas masks were inverse normalized to single subject space).

**EVC.** EVC was selected for each subject. Based on anatomical landmarks (visualized on the standard WFU PickAtlas T1-weighted image; <http://www.fmri.wfubmc.edu/cms/software>, in MNI space), we centered a sphere of 20 mm radius on the posterior calcarine sulcus (MNI coordinates 0, -100, -2) encompassing bilateral calcarine sulci and adjacent occipital cortex. Then, we mapped this sphere into the native space of each subject using the inverse normalization transformation. The transformation matrix was obtained with the SPM2 normalization function (source: MNI whole-brain EPI, target space: subject whole-brain EPI). For each participant, we selected visually activated voxels ( $p < 10^{-4}$ , uncorrected) within this sphere defined as those voxels significantly activated by face and vehicle stimuli when presented either above or below fixation during the experimental runs. These voxels were selected for the analysis of the experimental data.

**LO.** For each subject, we generated a  $t$  contrast “objects > scrambled objects.” The standard definitions of FFA (Kanwisher et al., 1997) and LOC (Malach et al., 1995) allow the possibility that the fusiform portion of LOC overlaps with the functionally defined FFA. To preclude this possibility, following the usage of prior fMRI studies that distinguished between subregions of LOC (Grill-Spector et al., 2001; Eger et al., 2008a, 2008b), we subdivided LOC into a posterior (LO) and an anterior portion (FUS) based on anatomical masks (WFU PickAtlas). Object-sensitive voxels ( $p < 10^{-4}$ , uncorrected) located either on the inferior or medial occipital gyri were allocated to LO and selected for the analysis of the experimental data.

#### Analysis of the main experiment

For each subject, we modeled the cortical response to experimental conditions for each of the five experimental runs separately on unsmoothed data. To better suit the goal of investigating fine-grained representations of object orientation in ventral visual cortex, we collapsed exemplar information during GLM estimation, thus specifying 20 experimental conditions of interest (2 categories  $\times$  2 positions  $\times$  5 angles). The onsets of the respective mini-blocks were entered into the GLM as regressors of interest and convolved with an HRF. This procedure yielded 20 parameter estimates per run representing the responsiveness of each voxel to the experimental conditions. Percentage signal changes in each ROI were computed with respect to the implicit baseline of the GLM and averaged across runs for each subject. Univariate statistical analyses were performed by means of repeated-measures ANOVAs with SPSS18. All reported  $p$  values were Greenhouse-Geisser corrected.

We also computed the same GLM with data that were smoothed with a Gaussian kernel (FWHM = 5 mm; GLM<sub>Smoothed</sub>). This was done to visualize the spatial distribution of univariate voxel sensitivities to stimulus position along the ventral stream and for selection of visually responsive voxels in early visual cortex. We independently assessed the contrasts [(faces & vehicles) below fixation > implicit baseline], and [(faces & vehicles) above fixation > implicit baseline]. To visualize the spatial distribution of voxel responsiveness to stimuli shown above, below, and both above and below fixation, we specified three sets of voxels ( $p < 10^{-5}$ , uncorrected): (1) voxels activated by stimuli below fixation only, (2) voxels activated by stimuli above fixation only, and (3) voxels activated both by stimuli above and below fixation. We then computed for each ROI a measure of voxel responsiveness to stimuli in both retinal positions, namely, the proportion of voxels significantly responding to stimuli presented in both retinal positions:  $PRV_{ROI} = (\text{above} \& \text{below}) / (\text{above} + \text{below})$ .

#### Pattern classification

Experimental data were subjected to three distinct multivoxel pattern classification analyses (Haxby et al., 2001; Carlson et al., 2003; Cox and Savoy, 2003; Haynes and Rees, 2005; Kamitani and Tong, 2005; Kriegeskorte et al., 2006) using linear support vector classifiers (SVCs). All analyses followed the same basic approach and were conducted separately for each ROI. In detail, for each subject, run, and ROI, we extracted parameter estimates associated with the experimental conditions under investigation. The parameter estimates for one ROI constitute feature vectors that represents the spatial response patterns associated with each condition. First, feature vectors from 4 of 5 runs were assigned to a training

dataset that was used to train a linear SVC with the regularization parameter set to  $C = 1$  in the LibSVM implementation (<http://www.csie.ntu.edu.tw/~cjlin/libsvm>). The trained SVC was used to classify feature vectors from the independent test dataset consisting of the fifth run. This was repeated five times (fivefold cross-validation), each time with feature vectors from a different run assigned to the independent test dataset. Decoding accuracies were averaged over these five iterations. Then we conducted second-level analyses on decoding accuracies across subjects by means of repeated-measures ANOVAs, and one-sample  $t$  tests against classification chance level (for pairwise classification, this was 50% classification accuracy).

#### Decoding Analysis I: classification of angle within category and position

In the first multivoxel pattern classification analysis, we determined separately for each category (i.e., faces and vehicles) whether activity patterns in EVC, LO, and rFFA allow decoding of angle using training and test data from the same position. That is, we tested whether EVC, LO, and rFFA contain position-dependent angle information. For this purpose, for every pairwise combination of presentation angles, using data from four runs, we trained a classifier to differentiate activity patterns evoked by each angle-pair when presented in a specific position (viz. either above fixation or below fixation). This procedure was iterated for each of the 10 viable angle-pair combinations and repeated for each of the two possible train-test position assignments: (1) train and test above fixation and (2) train and test below fixation. Finally, for the purpose of second-level statistical inference, classification accuracies were averaged, yielding a single mean classification value per subject.

#### Decoding Analysis II: classification of angle within category and across position

In the second pattern classification analysis, we determined whether activity patterns allow the read-out of angle information when training and test data were taken from different positions. This tests whether EVC, LO, and rFFA contain translation-tolerant angle information. The general procedure was identical to the previous classification analysis (see above). The difference was that, while in the former analysis training and testing were performed on patterns evoked by stimuli presented in the same position of the visual field, now training and testing were performed on patterns evoked by stimuli presented in different positions of the visual field. This difference is crucial and allowed us to evaluate whether angle information found in patterns of activation evoked by stimuli presented in one position of the visual field (e.g., above fixation) generalizes to patterns evoked by stimuli presented in a different position of the visual field (e.g., below fixation). The standard cross-validation procedure was conducted for all angle-pairs and possible position assignments to the training and test set. Again, for the purpose of second-level statistical inference, decoding accuracies were averaged, yielding a single mean classification value per subject.

#### Decoding Analysis III: classification of face and vehicle exemplar information

In an additional pattern analysis, we attempted to classify position-dependent and translation-tolerant exemplar information in face-sensitive areas rOFA, rSTS, and rFFA. First, we computed an additional GLM consisting of 40 conditions (2 exemplars  $\times$  2 categories  $\times$  2 positions  $\times$  5 angles) and conducted within-orientation classification of face and vehicle exemplar information. Analyses were performed both within and across positions, as described for the orientation decoding analyses above. For the purpose of second-level statistical inference, decoding accuracies within and across positions were separately averaged.

#### Comparison of the predictions of two representational models

To evaluate the goodness of fit for the two conceptual models (a monotonic representation of angle and a mirror-symmetric representation of angle) and the data obtained from rFFA, as well as for the purpose of model comparison, we performed a variant of Representational Similarity Analysis (RSA) (Kriegeskorte et al., 2008; Kriegeskorte, 2009). Our procedure consisted of four main steps: (1) translation of each conceptual model into a model similarity matrix (mSM), which serves as a

**Table 1. Model parameters<sup>a</sup>**

Parameters	Description
<b>Fixed parameters</b>	
$n_{Vox}$	Number of voxels ( $n = 120$ ). Mean number of voxels in rFFA across subjects, as estimated from our independent localizer run.
$n_{View}$	Number of view-tuning centers ( $n = 8$ ). Chosen to homogeneously span the whole range of possible orientations ( $360^\circ$ ).
<b>Free parameters</b>	
$\sigma$	Neural population tuning widths ( $\sigma = 20 - 60^\circ$ , step $10^\circ$ ). Corresponding to the range reported for face view-tuned neurons in primates (Perrett et al., 1991, 1998; Logothetis et al., 1995) and used in a recent related model of fMRI adaptation data (Andresen et al., 2009).
$n_{Patch}$	Number of view-tuned clusters assumed to reside within a voxel. We broadly constrained this parameter based on previous studies of face orientation encoding in primate IT cortex. We tested clustering levels ranging from approximately that of a macrocolumn (radius = $250 \mu\text{m} \rightarrow m = 2^4 = 16 \text{ c/v}$ ) to a quarter of a macrocolumn (radius = $62.5 \mu\text{m}$ , $m = 2^8 = 256 \text{ c/v}$ ). Step of powers of two ( $2^n$ , $n = 4, n = 5, n = 6, n = 7, n = 8$ ). $\text{c/v}$ , Clusters per voxel.
$K$	Modulation parameter of the distribution $Q_s$ assigning probabilities to the multiple orientation-tuned populations. Range $-1, 1$ ; step $0.2$ . Precisely, $K$ was used to modulate a cardioid function (Jeffreys, 1961) by means of a multiplication operation. Refer to Equations 4 and 5.

<sup>a</sup>Listed symbols (in italics) denote model parameters used in the simulation of empirical similarity matrices associated with head rotation in-depth in the FFA. Each symbol is accompanied by a brief description of the corresponding parameter and its exact value, or range of values, as appropriate.

“model template”; (2) the computation of an empirical similarity matrix (eSM) for each subject reflecting the pairwise linear correlations (Pearson’s  $r$ ) of the brain activity patterns associated with a set of experimental conditions of interest; (3) the assessment of the fit of the model to the data by means of the correlation (Spearman’s  $\rho$ ) between each eSM and the model template (i.e., mSM); and (4) the statistical test of the relevant null hypothesis by means of a permutation test.

Because we are interested in the translation-invariant aspect of the representation of orientation information in rFFA, relevant empirical SMs are those describing for each object category (faces and vehicles) similarity relationships of fMRI pattern vectors related to orientation ( $-90^\circ, -45^\circ, 0^\circ, 45^\circ, 90^\circ$ ) between positions (i.e., orientation below fixation to orientation above fixation) (see Fig. 4A). For each subject and object category, eSMs were computed as follows. First, we estimated the mean response for each experimental condition by averaging GLM parameter estimates across runs. Then, we computed for each subject two SMs, one per object category. Correlations were Fisher- $Z$  transformed for all subsequent analyses. In a final step, and with the goal of evaluating the statistical significance of the representational correlation between a model and the empirical data, we tested the null hypothesis ( $H_0$ ) of no positive correlation (Spearman’s  $\rho$ ) between each model template and the data. For this purpose, we used an exact sign permutation test (Good, 2004). Similarly, to evaluate whether for a specific object category one of the two models better fits the data, we tested the null hypothesis of zero difference between the representational correlations associated with each model across the subject population. Only for plotting purposes, bootstrap procedures (Efron and Tibshirani, 1994) were used to estimate the sampling distribution of the population correlation between each model and the data. A total of 10,000 resamples with replacement were drawn from our subject pool. For each resample, an average SM was calculated and the correlation with each model template computed. Estimated distributions are shown in Figure 4.

### Computational modeling

The section is organized as follows: Model overview and formulas, Sampling of the cortical surface by fMRI voxels, Signal-to-noise equalization, and Model extension: probing for mirror-symmetric tuning. Model parameters are summarized in Table 1.

### Model overview and formulas

To assess a biologically motivated interpretation of our main classification results in FFA, we modeled the eSMs associated with our face stimulus between positions. As recently emphasized by Kriegeskorte et al. (2008), SMs provide compact characterizations of the representational structure in a given brain or model. Thus, the adequacy of a model family can be assessed based on the fit between the SM predicted by the model and the SMs obtained from experimental measurements.

A wealth of neuroscientific research has provided evidence of neurons selectively tuned to faces in primate ventral visual cortex. These neurons usually display a unimodal tuning curve with a preferred orientation, are tolerant to translation within the visual field, and have been reported to be clustered at spatial scales approximately consistent with that of a mac-

rocolumn (Fujita et al., 1992; Wang et al., 1996). Importantly, the response selectivity of such view-tuned neural clusters has been shown to be remarkably tolerant to retinal translation (Wang et al., 1998). Concerning the neural representation of different face views, both monkey electrophysiology (Hasselmo et al., 1989; Perrett et al., 1998) and human fMRI data (Tong et al., 2000; Yue et al., 2011; Axelrod and Yovel, 2012) suggest that the frontal views of faces might be overrepresented with respect to the lateral and posterior views of faces. Based on these observations, we argue that inhomogeneously distributed clusters of orientation-tuned neurons may account for the representational structure to face orientation observed in rFFA. To substantiate this claim, we proceeded to simulate distributed fMRI response patterns to our experimental conditions. A set of model templates (mSMs) was computed, where each mSM was estimated as the expectation value of a specific model parameter combination, thus allowing us to evaluate the fit of each mSM to the empirical data, and consequently make inferences concerning the underlying neural representations based on the parameter values of the model fit.

To assess whether mixtures of translation-invariant view-tuned neural populations (Wang et al., 1996, 1998) can explain the empirically observed representation of angle across position, we model the BOLD response in each voxel as the summed responses of view-tuned neural populations ranging from front view ( $0^\circ$ ) to back view ( $180^\circ$ ) in both directions of rotation. The response of each view-tuned population  $s = (1 \dots n_{View})$ , with  $n_{View} = 8$  (see Fig. 6, top left), to a stimulus presented at an angle  $\theta$  was modeled by a Gaussian, as follows:

$$Res_s(\theta_{Stimulus}) = G(\theta_{Stimulus} - \mu_s, \sigma), \quad (1)$$

centered at a preferred view  $\mu_s$  with a tuning width  $\sigma$ , where

$$\mu_s = \frac{s - 4}{n_{View}} \cdot 360^\circ. \quad (2)$$

Thus, the total response of a voxel  $k$  to a stimulus presented to the observer is the sum of the view-tuned population responses, as follows:

$$\beta_k(\theta_{Stimulus}) = \frac{1}{n_{View}} \sum_s w_{sk} Res_s(\theta_{Stimulus}) \quad (3)$$

where  $w_{sk}$  denotes the contribution of each view-tuned neural population  $s$  to voxel  $k$ .

We treat each voxel as a random sample from a pool of view-tuned neural clusters, which is characterized by the probability  $Q_s$  of a neural cluster to be tuned to a particular view. We modeled this probability by a cardioid function (Jeffreys, 1961), as follows:

$$Q_s = \frac{1}{n_{View}} (1 + K \cos \mu_s), \quad (4)$$

where the parameter  $K$  serves to specify the degree of variation in the probabilities assigned to each view-tuned population  $s$  (see Fig. 6B). Thus, depending on  $K$ , the possible distributions of  $Q_s$  can range from a

homogeneous distribution ( $K = 0$ , such that every view-tuned population is equally represented) to inhomogeneous distributions where frontal views ( $K > 0$ ) or back views ( $K < 0$ ) are overrepresented. We chose a cardioid function because it is the simplest periodic function able to model an overrepresentation of the frontal views, as suggested by prior findings from monkey electrophysiology (Hasselmo et al., 1989; Perrett et al., 1998) and human fMRI (Tong et al., 2000; Yue et al., 2011; Axelrod and Yovel, 2012), and the univariate response profile observed in our own data.

Each voxel is a sample from such a pool of view-tuned neural clusters, with a granularity defined by the typical number  $n_{Patch}$  of view-tuned clusters assumed to reside within a voxel, which leads to a multinomial distribution, as follows:

$$\vec{W} \approx M(n_{Patch}, \vec{Q}) \quad (5.1)$$

where

$$\vec{Q} = (Q_1, Q_2, \dots, Q_s) \quad (5.2)$$

$$\vec{W} = (W_1, W_2, \dots, W_s) \quad (5.3)$$

Based on this approach, we modeled the empirically observed multivoxel response patterns to the experimental conditions of interest as a mixture of signal and noise, as follows:

$$\vec{\beta} = (\beta_1, \beta_2, \dots, \beta_{Condition}) \quad (6)$$

$$\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_{Condition}) \quad (7)$$

where  $\beta_{Condition}$  denotes the signal (i.e., the expected multivoxel pattern response for a particular experimental condition), whereas the corresponding vector  $\beta$  groups the expected multivoxel pattern responses for all relevant experimental conditions. Similarly,  $\xi_{Condition}$  denotes the noise component for a particular experimental condition, whereas the corresponding vector  $\xi$  groups the expected noise components for all the relevant experimental conditions. Thus, for a particular simulation, the mixture of signal plus noise for a family of conditions (for details, see Signal-to-noise equalization below) will be obtained by the following:

$$\vec{\beta}' = \vec{\beta} + \vec{\xi} \quad (8)$$

For each concrete combination of free parameters of the model, we ran 1000 simulations, then computed for each simulation a SM, and averaged across SMs to obtain a model template (i.e., mSM) for each combination of free parameters of the model. The repetition of this procedure for every combination of parameters resulted in a set of model templates spanning our parameter space. In a final step, the goodness of fit between mSMs and empirical SMs was evaluated by means of correlations (Spearman's  $\rho$ ), in line with previous publications (Kriegeskorte et al., 2008; Kriegeskorte, 2009).

#### Sampling of the cortical surface by fMRI voxels

fMRI normally samples the cortex with a regular voxel grid, whereas the cortical mantle ( $\sim 2$  mm thick) intricately folds forming sulci and gyri. Accordingly, the proportion of gray matter sampled by a voxel is certainly not constant. The maximal response of a voxel must depend on the proportion of gray matter it samples. As a plausible approximation, we specified the gray matter proportion sampling distribution of our voxels by means of a transform,

$$g(x) = 3x^2 - 2x^3, \quad (9)$$

where  $x$  is uniformly distributed on  $[0, 1]$ . This way we modeled the probability distribution of obtaining a given proportion of gray matter in the idealized case where a 2 mm flattened gray matter sheet is sampled by a sphere of diameter 2 mm. We assumed that all points in space are sampled with equal probability with proportions ranging from 0% to 100%. Based on this approach, for each particular simulation, we obtained a weighting vector. Last, gray matter proportions for each voxel were used

to scale simulated response amplitudes to the experimental conditions of interest.

#### Signal-to-noise equalization

A central aspect our model aims to capture is the effect of signal-to-noise differences among experimental conditions on the corresponding dissimilarity matrices while respecting the mean signal-to-noise levels empirically observed. Accordingly, for every set of parameters, we equalized the mean signal-to-noise ratios (SNRs) of the modeled response patterns to the mean SNR estimated from the data. Data and model SNRs were computed in exactly the same manner. We estimated the subject population SNR as the mean SNR across subjects. The SNR for each subject was in turn defined as the mean across the experimental conditions of interest (i.e., faces in various orientations). For modeling purposes, voxel-noise was assumed to be i.i.d. according to  $N(0, \sigma_{Noise}^2)$ . SNR was equalized between simulations and data by means of an iterative algorithm accepting values within a tight margin of the specified target SNR ( $SNR_{DATA} \pm 0.0001$ ).

#### Model extension: probing for mirror-symmetric tuning

To assess whether incorporating various degrees of mirror-symmetric tuning to the model may better explain the empirically observed representation of head angle, thereby providing evidence of mirror-symmetric tuning in rFFA, we modeled the BOLD response in each voxel as the summed responses of neural populations ranging in their degree of mirror-symmetric tuning from 0% to 100%. In detail, for all view-tuned patches associated with a particular simulation, we modified the default unimodal Gaussian tuning function to head orientation by including a second mode, also Gaussian, with a maximum response level specified as a percentage of the maximal response of the unimodal (basis) Gaussian. This percentage was controlled by an additional parameter tau ( $\tau$ ; range: 0%–100%, step 10%). Thus,  $\tau$  was used to specify for each mSM the degree of mirror-symmetric tuning associated with that simulation. Because mirror-symmetrically placed Gaussian tuning functions can substantially overlap if tuning-width is broad relative to the angular distance from the frontal view of the preferred orientation, we specified tuning functions in points of overlap as the maximum among the two Gaussians (thus always prescribing maximal response levels for a patch to its true preferred orientation). The expectation of each admissible parameter combination was obtained as above. Finally, the utility of including the additional parameter  $\tau$  (mirror symmetry) was evaluated by means of a leave-one-subject-out cross-validated model-comparison approach. Thus, we directly compared models (1) some degree of mirror symmetry versus (2) no mirror symmetry (i.e., not including parameter  $\tau$ ). For each cross-validation fold, Models 1 and 2 were computed as the mean of the fitted parameters across training subjects. The goodness of fit of the estimated models was compared using the eSM of the left-out subject. In each fold, the model exhibiting the highest correlation (Spearman) to the left-out subject's eSM received one vote. A model receiving a majority of votes should be considered as preferable. If the more complex of two models does not prove to be preferable, the simpler model should be favored on grounds of parsimony.

## Results

Participants performed a brightness discrimination task orthogonal to the experimental conditions of interest: object category (face, vehicle), orientation in-depth ( $-90^\circ, -45^\circ, 0^\circ, 45^\circ, 90^\circ$ ), and retinal location ( $3^\circ$  above or below fixation) (Fig. 1). We examined the associated multivariate, unsmoothed fMRI patterns in prespecified ROIs measured from human occipitotemporal cortex. Specifically, we assessed the discriminability and position dependence of orientation information in three critical processing stages along the ventral processing stream: early visual areas, LO, and FFA.

#### Univariate analyses

First, to characterize the mean responses associated with our experimental manipulations in each ROI, we subjected the data to

univariate analyses. Results are summarized in Figure 2B and the outcome of statistical analyses presented below.

### EVC

The observed mean responses are on average comparable across categories (faces and vehicles). A three-way ANOVA, with category, position, and orientation (i.e., angle of presentation of a stimulus) as repeated measures, revealed no significant main effect of category ( $F_{(1,7)} = 0.78, p = 0.41$ ), position ( $F_{(1,7)} = 0.001, p = 0.97$ ), or orientation ( $F_{(4,28)} = 2.32, p = 0.13$ ). The only significant effect observed was an interaction of orientation by category ( $F_{(4,28)} = 4.15, p = 0.027$ ). The latter observation is in agreement with the known retinotopic organization of early visual areas (Hansen et al., 2004), such that, for example, profile views of vehicles are more elongated along the horizontal axis than faces, and this property is similarly reflected in the mean response profiles of each retinal location.

### LO

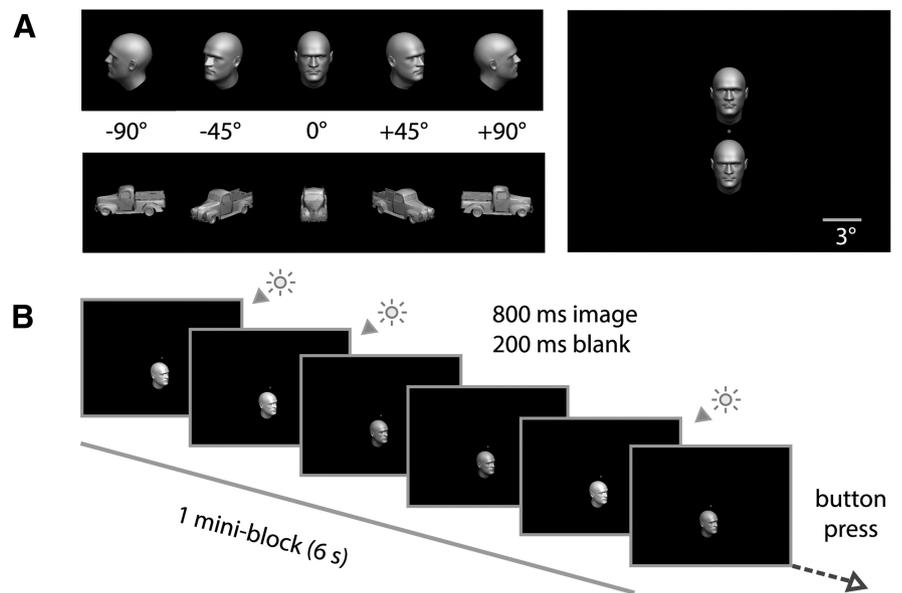
Three-way ANOVA with category, position, and orientation as repeated measures revealed no main effects of category ( $F_{(1,7)} = 1.26, p = 0.30$ ) or orientation ( $F_{(4,28)} = 0.51, p = 0.59$ ). However, we did observe a significant main effect of position ( $F_{(1,7)} = 5.73, p = 0.048$ ), and a significant interaction of category by position ( $F_{(1,7)} = 37.52, p < 0.001$ ). The latter effects may be explained by a reported lower visual field response bias in LO (Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008). In addition, we found a significant interaction of orientation by category ( $F_{(4,28)} = 3.78, p = 0.031$ ), consistent with our observations in EVC and reported evidence of weak retinotopy in LO (Larsson and Heeger, 2006; Sayres and Grill-Spector, 2008).

### rFFA

As is expected of a face-selective cortical region, stronger responses to faces than vehicles were observed regardless of the angle of presentation (Fig. 2B, right). A three-way ANOVA with category, position, and orientation as repeated measures confirmed this observation. We observed a highly significant main effect of category ( $F_{(1,7)} = 116.33, p < 0.001$ ), as well as a significant effect of angle ( $F_{(4,28)} = 4.94, p = 0.021$ ). No main effect of position ( $F_{(1,7)} = 0.17, p = 0.69$ ) or interaction effects were found (all  $F$  values  $< 1.41$ ). In line with previous studies addressing the representation of face orientation in FFA (Yue et al., 2011; Axelrod and Yovel, 2012), subjecting our data for faces to two-way ANOVA with position and orientation as repeated measures revealed a significant effect of orientation ( $F_{(4,28)} = 3.90, p = 0.044$ ), which was accompanied by a significant quadratic trend ( $F_{(1,7)} = 9.02, p = 0.02$ ), thus confirming recent evidence of a stronger representation of frontal than lateral views of faces (Yue et al., 2011; Axelrod and Yovel, 2012).

### Position dependence of voxel responses

To visualize the spatial distribution of voxel sensitivities to stimulus location in ventral visual cortex, in Figure 2C, we show for a representative subject those voxels significantly responding only to stimuli presented above fixation in green, only to stimuli below



**Figure 1.** Experimental design. **A**, Images of faces and vehicles were shown at five different rotational angles (orientations), either 3° above or below fixation. **B**, Subjects were instructed to maintain fixation at a central spot and perform a brightness detection task on objects presented within a miniblock.

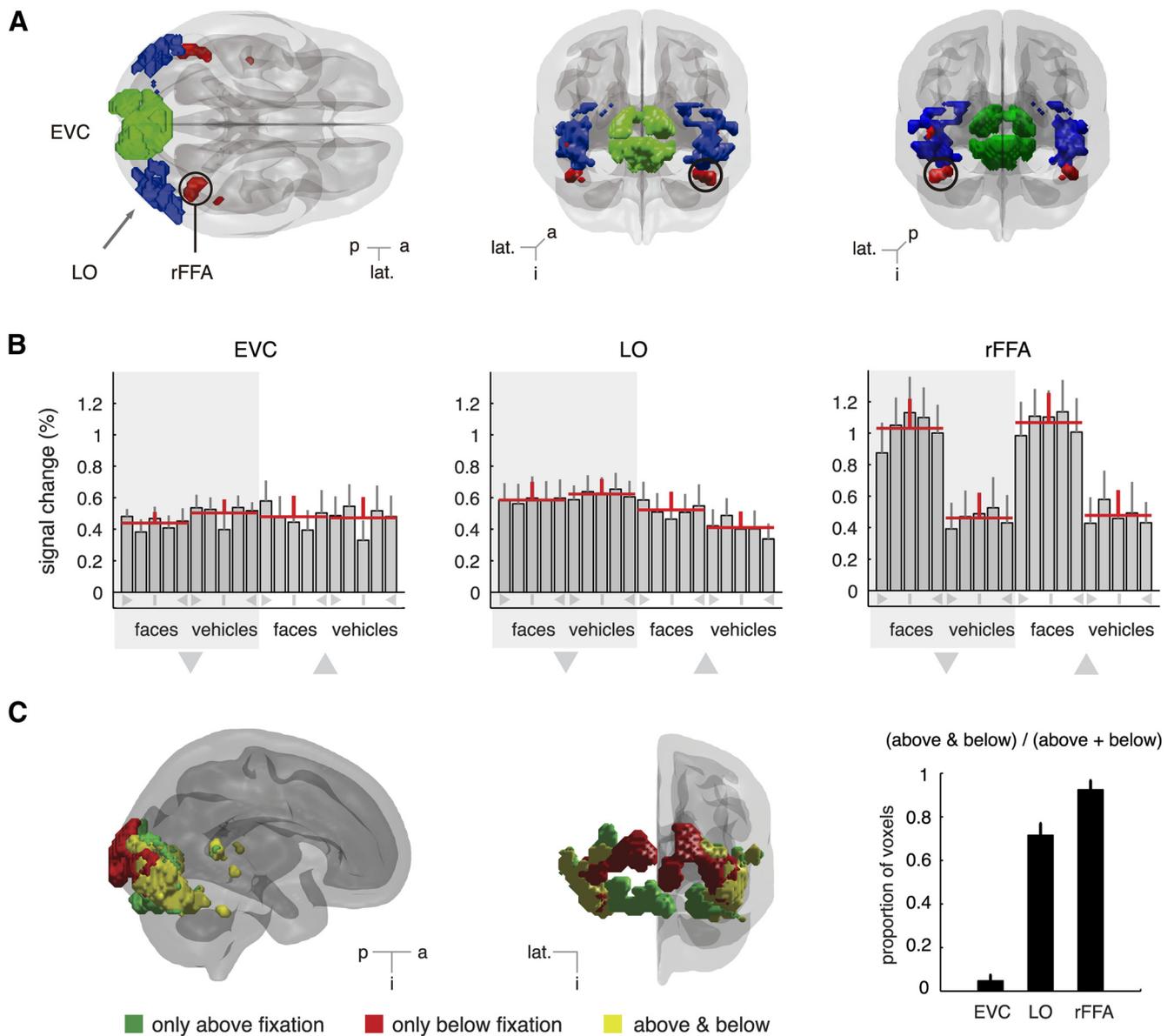
fixation in red, and voxels responsive to both locations in yellow. Summary statistics across the population are summarized in Figure 2C (right). As expected, voxels in the most posterior portions of the brain proved position dependent, with dorsal and ventral voxels responding, respectively, to lower and upper visual field stimulation. On the other hand, voxels sensitive to both locations were predominant in more anterior portions of the ventral stream, including lateral occipital areas. The first cortical area analyzed here where voxels consistently responded to both locations of the visual field coincides with LO, consistent with evidence of translation-tolerant object representations at this level of the visual hierarchy (Grill-Spector and Malach, 2004).

### SVM classification analyses

To learn whether orientation information is encoded in the fine-grained patterns measured from EVC, LO, and rFFA, we conducted classification via a linear support vector machine (SVM) (Vapnik, 1995) on the fine-grained (i.e., unsmoothed) fMRI activation patterns from these ROIs (Cox and Savoy, 2003; Haynes and Rees, 2006). By “activation pattern” we mean the set of run-wise GLM parameter estimates in an ROI corresponding to an experimental condition of interest. Given that the position dependence of multivoxel representations at the level of categories has been recently investigated (Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008; Carlson et al., 2011; Cichy et al., 2011, 2013), we focused on the more subtle encoding of orientation information within categories.

### LO: shape encoding partially generalizes across locations

We observed significant orientation information in LO regardless of object category. In agreement with previous studies (Larsson and Heeger, 2006; Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008; Carlson et al., 2011; Cichy et al., 2013), we found in LO a significant main effect of train-test position correspondence (from here on referred to as “position”) on orientation classification ( $F_{(1,7)} = 9.64, p = 0.017$ ) (Fig. 2B). We found no effect of category ( $F_{(1,7)} = 0.68, p = 0.44$ ) or interaction of position by category ( $F_{(1,7)} = 0.57, p = 0.48$ ). One-sample  $t$



**Figure 2.** ROIs and univariate analyses. **A**, EVC, LO, and inside a circle the rFFA for a representative subject are shown from the superior (left) and posterior (middle) and anterior (right) views. **B**, Percentage signal changes. Red horizontal lines indicate the mean activation level across subjects and views for each combination of object category and position. Large triangles below the shaded areas represent conditions corresponding to stimuli shown below or above fixation. Symbols below bars represent orientation: vertical line indicates 0°; triangles represent ±90°; not labeled represents ±45°. Whereas EVC and LO show comparable response levels for faces and vehicles shown above or below fixation, rFFA exhibits an evident preference for faces over vehicles. An overrepresentation of the frontal views of faces can also be noted. **C**, Responsiveness in ventral visual cortex to stimuli shown above and below fixation. 3D rendering of significantly responsive voxels in a representative subject ( $p < 10^{-5}$ , uncorrected; see Methods for details). Early visual areas exhibit a strong position dependence, that is, voxels in posterior occipital areas respond to stimuli presented either above fixation (ventral voxels, green) or below fixation (dorsal voxels, red), whereas voxels responsive to both locations (yellow) predominate in lateral–occipital and ventral areas. Consistency of these observations across subjects in our ROIs is summarized in the contiguous barplot. Error bars indicate SEM. p, Posterior; a, anterior; lat., lateral; i, inferior.

tests confirmed the significance of position-dependent information for both faces ( $t_{(7)} = 4.41, p = 0.003$ ) and vehicles ( $t_{(7)} = 7.06, p < 0.001$ ), whereas additional paired  $t$  tests revealed more position-dependent than position-independent information (faces:  $t_{(7)} = 2.46, p = 0.044$ ; vehicles:  $t_{(7)} = 3.41, p = 0.011$ ). However, position-independent information was present for both faces ( $t_{(7)} = 4.56, p = 0.003$ ) and vehicles ( $t_{(7)} = 3.26, p = 0.014$ ). Importantly, this pattern of results was not observed in EVC (Fig. 3). This finding supplements reports of translation-tolerant encoding of categorical (Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008; Carlson et al., 2011; Cichy et al., 2011, 2013) and subcategorical (Cichy et al., 2011) object information in ventral visual cortex. It also shows that, at this level of the visual

hierarchy, encoding of object shape is category-general and translation-tolerant (Grill-Spector et al., 1999).

**FFA: category-selective encoding of orientation**

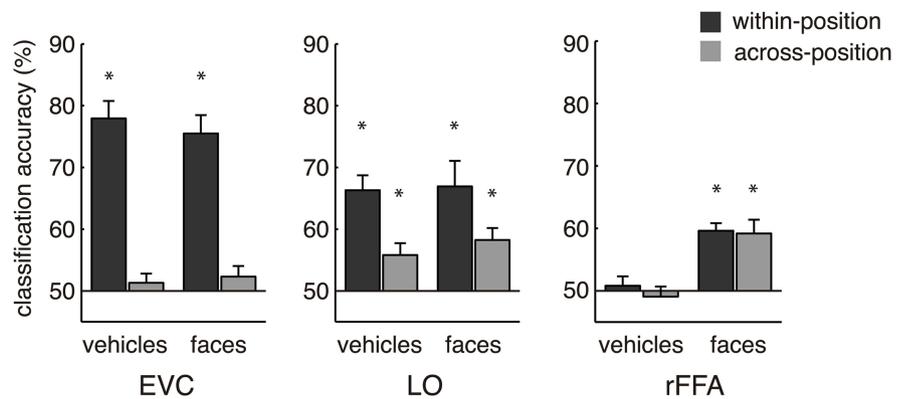
We then investigated whether the availability of orientation information in FFA depends on object category. We found orientation information in rFFA, but only for faces, as indicated by a highly significant main effect of object category on the classification of orientation (two-way repeated-measures ANOVA;  $F_{(1,7)} = 42.80, p < 0.001$ ). Interestingly, we observed no main effect of position ( $F_{(1,7)} = 0.66, p = 0.44$ ), suggesting a category-selective and translation-invariant code (Fig. 3). We found no evidence of an interaction ( $F_{(1,7)} = 0.22, p = 0.65$ ) or evidence of orientation

information for vehicles either within position ( $t_{(7)}, t = 0.57, p = 0.58$ ) or across position ( $t_{(7)}, t = 0.62, p = 0.57$ ). In stark contrast to vehicles, one-sample  $t$  tests for faces revealed significant position-dependent ( $t_{(7)}, t = 8.56, p < 0.001$ ) and position-independent information ( $t_{(7)}, t = 4.46, p = 0.003$ ). A direct comparison of faces and vehicles by additional paired  $t$  tests confirmed significantly more face than vehicle information (within position:  $t_{(7)} = 8.12, p < 0.001$ ; across position:  $t_{(7)} = 3.86, p = 0.006$ ). Together, these findings suggest that rFFA may implement a category-selective and translation-invariant representation of face orientation.

Then, we investigated the availability of orientation information in additional face-preferring areas rOFA and rSTS. Two-way repeated-measures ANOVA in these areas revealed no main effects of category (rOFA:  $F_{(1,6)} = 0.2, p = 0.67$ ; rSTS:  $F_{(1,7)} = 0.62, p = 0.46$ ) or position (rOFA:  $F_{(1,6)} = 0.54, p = 0.49$ ; rSTS:  $F_{(1,7)} = 5.47, p = 0.052$ ). Further, we observed no evidence of category-selective encoding of orientation information, as indicated by nonsignificant interaction effects of orientation by category (rOFA:  $F_{(1,6)} = 0.31, p = 0.60$ ; rSTS:  $F_{(1,7)} = 0.91, p = 0.37$ ). Statistical testing of mean decoding accuracies against chance level, except for vehicles in rSTS across position (mean = 53.4%,  $t_{(7)} = 3.02, p = 0.02$ ) did not provide significant evidence of position-dependent or translation-tolerant encoding of orientation information in rOFA (vehicles: within position 51.6%, across position 54.1%; faces: within position 53.9%, across position 53.4%;  $t$  tests, all  $p > 0.08$ ) or in rSTS (vehicles: within position 49.8%; faces: within position 46.6%, across position 53.5%;  $t$  tests, all  $p > 0.07$ ). Although classification of orientation proved significant for vehicles across position in rSTS, it was not significantly different from classification within position (paired  $t$  test,  $t_{(7)} = 1.42, p = 0.2$ ) or from faces within or across position (paired  $t$  tests,  $t_{(7)} = 2.05, p = 0.08$ ;  $t_{(7)} = 0.02, p = 0.98$ ). In sum, classification of orientation in rOFA and rSTS was found to be unreliable and provided no significant evidence of category selectivity.

### Classification of exemplars in face-preferring areas

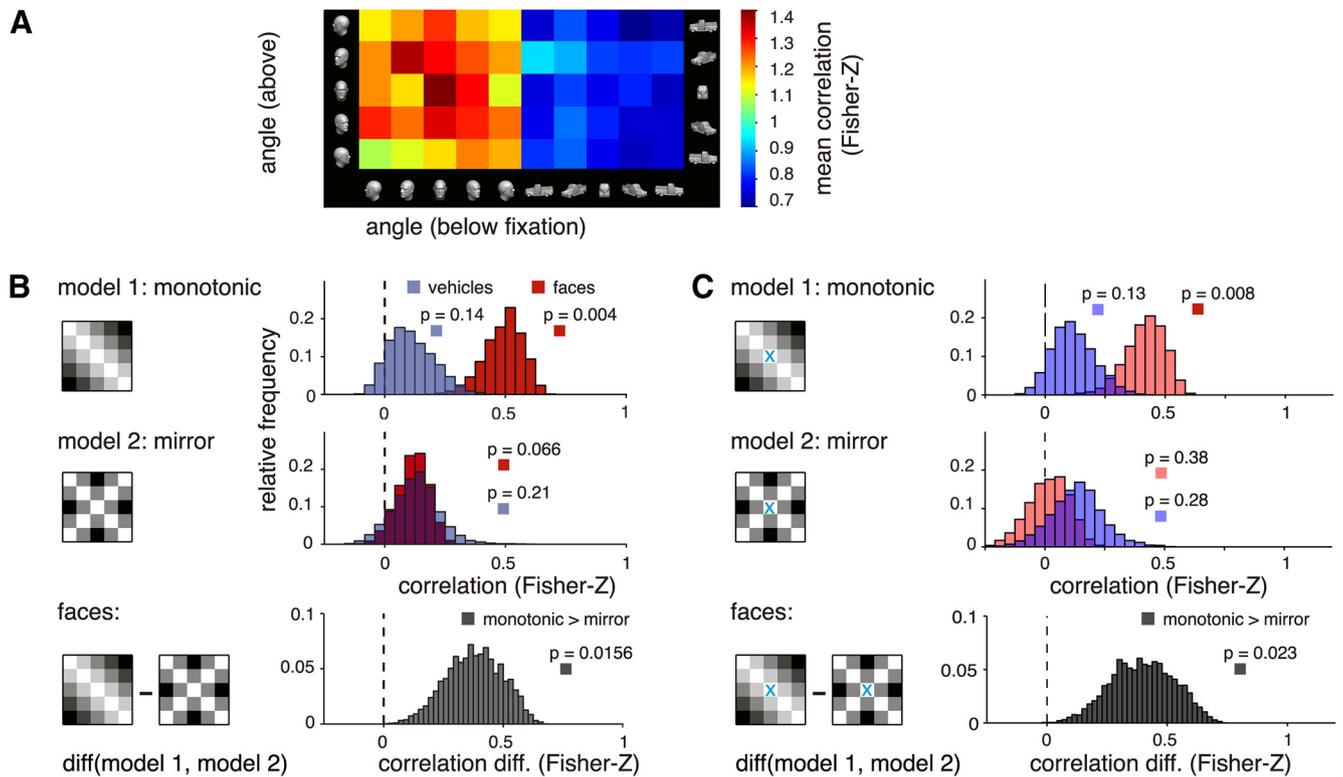
We then asked whether distributed responses in face-preferring areas rOFA, rSTS, and rFFA may reliably distinguish face and/or vehicle exemplar information. We found no evidence of position-dependent encoding of face or vehicle exemplar information in rOFA (faces: 53.3%;  $t_{(6)} = 2.0, p = 0.09$ ; vehicles: 55.1%,  $t_{(6)} = 2.2, p = 0.07$ ), rSTS (faces, 49%;  $t_{(7)} = -0.44, p = 0.67$ ; vehicles: 52.6%,  $t_{(7)} = 1.62, p = 0.16$ ), or rFFA (faces: 50.8%,  $t_{(7)} = 0.4, p = 0.7$ ; vehicles: 48.9%,  $t_{(7)} = -1.1, p = 0.32$ ). Likewise, we did not observe evidence of translation-tolerant exemplar encoding in rOFA (faces: 47.3%;  $t_{(6)} = -1.6, p = 0.16$ ; vehicles: 48.7%,  $t_{(6)} = -0.8, p = 0.46$ ), rSTS (faces: 52.9%;  $t_{(7)} = 1.2, p = 0.27$ ; vehicles: 48.3%,  $t_{(7)} = -0.77, p = 0.47$ ), or rFFA (faces: 48.9%,  $t_{(7)} = -1.08, p = 0.32$ ; vehicles: 47.8%,  $t_{(7)} = -1.05, p = 0.33$ ). The observed lack of exemplar information may be partially explained by the brightness detection task performed by subjects in our study (diverting attention from identity information), as well as the parafoveal locus of stimulation.



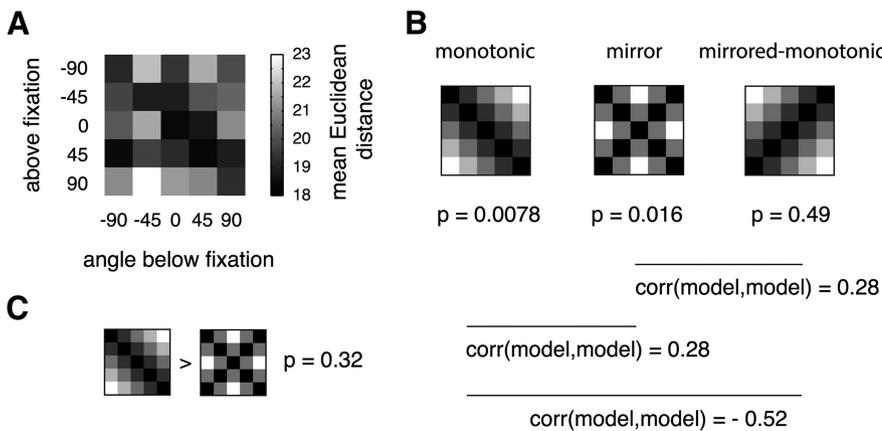
**Figure 3.** Decoding of rotational angle within each object category. Two analyses were performed. “Within position” refers to classifiers trained and tested on the same positions (i.e., above — above, or below — below), whereas “across position” refers to classifiers trained and tested in different positions. For each analysis, decoding accuracies were averaged across all possible angle pairs. Orientation classification in EVC was possible only within position, whereas higher visual areas exhibited considerable tolerance to translation. Whereas LO had information about rotational angle for both faces and vehicles, rFFA only encoded orientation information about faces. Asterisks indicate classification accuracies significantly different ( $p < 0.05$ ) from chance level (50%).

### Predominance of a monotonic code reflecting head angular disparity

Two basic representational schemes for face view information have been observed in primate face-selective areas (Freiwald and Tsao, 2010): (1) a monotonic code reflecting angular distance, meaning that to increasing angular distances between faces correspond increasingly dissimilar neural representations; and (2) a mirror-symmetric scheme, implying that to mirror-symmetric views (e.g., a left and a right profile) correspond similar neural representations. To assess whether response patterns in rFFA can be described by any one of these schemes, we used a variant of RSA (Kriegeskorte et al., 2008). To circumvent low-level effects that could confound the underlying representational structure, we focused our analyses on between position eSMs. Our main findings are summarized in Figure 4. Essentially, for faces we found a significantly better correlation of the monotonic model compared with the mirror-symmetric scheme (sign-permutation test,  $p = 0.0156$ ). Further testing confirmed a significant correlation between the eSMs collected for faces and the monotonic model (sign-permutation test,  $p = 0.004$ ), whereas evidence of such agreement with the mirror-symmetric scheme was lacking (sign-permutation test,  $p = 0.066$ ), at least when using a correlation-based RSA that is not sensitive to the mean signal (for a different analysis that is sensitive to mean differences between rotational angles, see below). We found no evidence of monotonic or mirror-symmetric encoding of vehicles (sign-permutation test,  $p = 0.14$  and  $p = 0.21$ , respectively). Moreover, according to the population correlation between the data for faces and the mirror-symmetric model, this model accounts for only 1.36% of the total variance (bootstrapped  $CI_{95\%} = 0\%, 4.8\%$ ), which is considerably less than the 24.3% of variance explained by the monotonic model ( $CI_{95\%} = 12.3\%, 37.1\%$ ). Additionally, we note that the nonsignificant trend for weak mirror symmetry observed for faces and the mirror-symmetric model may be due to the peak correlation value observed for the frontal views (Fig. 4A). This feature is uninformative regarding mirror-symmetric encoding of face orientation because it lies on the axis of symmetry. Therefore, we repeated our RSA analyses while omitting the cell corresponding to the frontal view [i.e., cell ( $0^\circ, 0^\circ$ )] from mSMs and eSMs (Fig. 4C). The outcome of this analysis shows that, when



**Figure 4.** Representational similarity analysis. **A**, Mean similarity matrices across subjects of multivariate fMRI responses in rFFA for different rotational angles, shown separately for faces and vehicles. **B**, We compared the empirical similarity matrices with two models: monotonic and mirror-symmetric (“mirror”) angle codes. We observed a significant correlation only for faces and only for the monotonic model ( $p$  values: exact sign-permutation tests). **C**, We compared the empirical similarity matrices with the same models used in **B**, but omitting cell  $(0^\circ, 0^\circ)$  from the analysis. This feature lies on the axis of symmetry and is therefore uninformative regarding mirror-symmetric encoding. The trend observed in **B** of a correlation between our data for faces and the mirror-symmetric model is no longer apparent when ignoring this feature.



**Figure 5.** Representational similarity analysis based on the Euclidean distance. **A**, Mean similarity matrix across subjects of multivariate fMRI responses in rFFA for different rotational angles for faces. **B**, We compared empirical similarity matrices with three models: monotonic model, mirror-symmetric model, and mirrored version of the monotonic model. Significant correlations with the data were observed for the monotonic and the mirror-symmetric models ( $p$  values: exact sign-permutation tests). The mirrored monotonic model was not significantly correlated with the data. Because the monotonic model is partially correlated to the mirror-symmetric model, conclusive evidence for mirror symmetry would require that the mirror-symmetric model outperforms the monotonic model. **C**, The mirror-symmetric model did not outperform the monotonic model.

ignoring this feature, the expected population correlation between our data for faces and the mirror-symmetric model is  $r = 0.017$  (bootstrapped  $CI_{95\%} = -0.13, 0.15$ ;  $p = 0.38$ , sign-permutation test). Importantly, agreement between our data and the monotonic model without cell  $(0^\circ, 0^\circ)$  remained significant ( $r = 0.42$ ; Bootstrapped  $CI_{95\%} = 0.27, 0.54$ ;  $p = 0.0078$ , sign-permutation test). In sum, these results demon-

strate that the representation of face orientation in rFFA is best accounted for by a monotonic population-code reflecting angular disparity.

We additionally performed an RSA using the Euclidean distance metric to assess the effects of a distance measure that is sensitive to mean response levels (Fig. 5). We found that, in this case, the monotonic model and the mirror-symmetric model were not distinguishable (Fig. 5). Both our study and a previous study (Axelrod and Yovel, 2012) find that the univariate response to frontal views is stronger than to side views. In these cases, using a distance measure sensitive to the mean is likely to be influenced by this source of mirror symmetry, albeit possibly indirectly. This, however, depends on whether the mean response in a region is considered a useful coding dimension (see Discussion).

**Computational modeling**

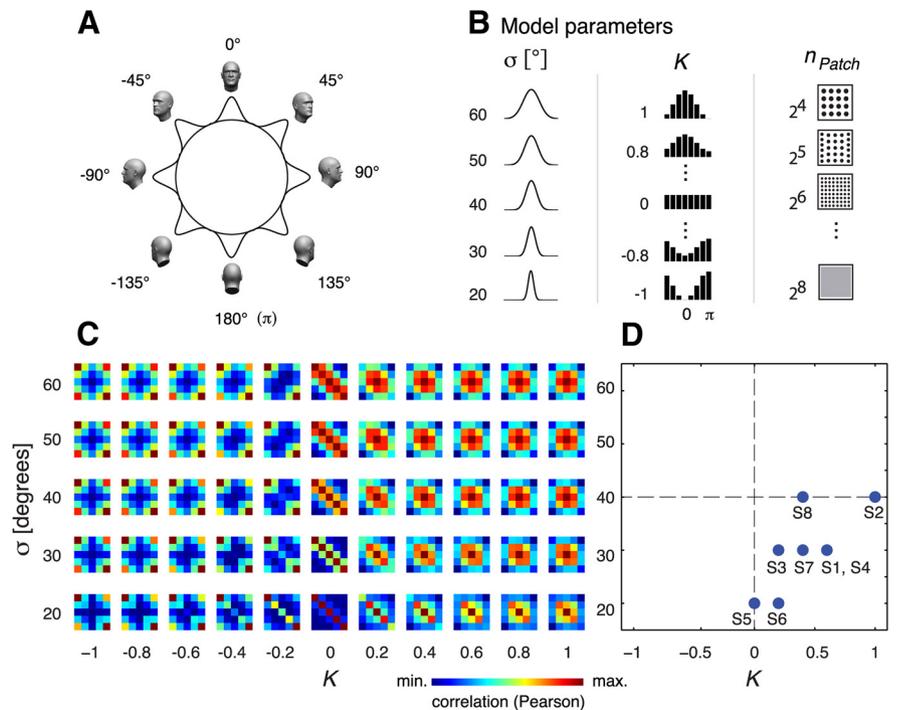
After ascertaining that the representational scheme best characterizing our data corresponds to a monotonic code reflecting head angular disparity, we considered which neurobiological properties might account for our fMRI measurements. Thus, we explored whether a family of computational models, based on empirical observations from optical imaging and electrophysiology in primates (Hasselmo et al., 1989;

Perrett et al., 1991, 1998; Fujita et al., 1992; Logothetis et al., 1995; Wang et al., 1996), might explain the empirically observed dissimilarity structure for faces rotated in-depth.

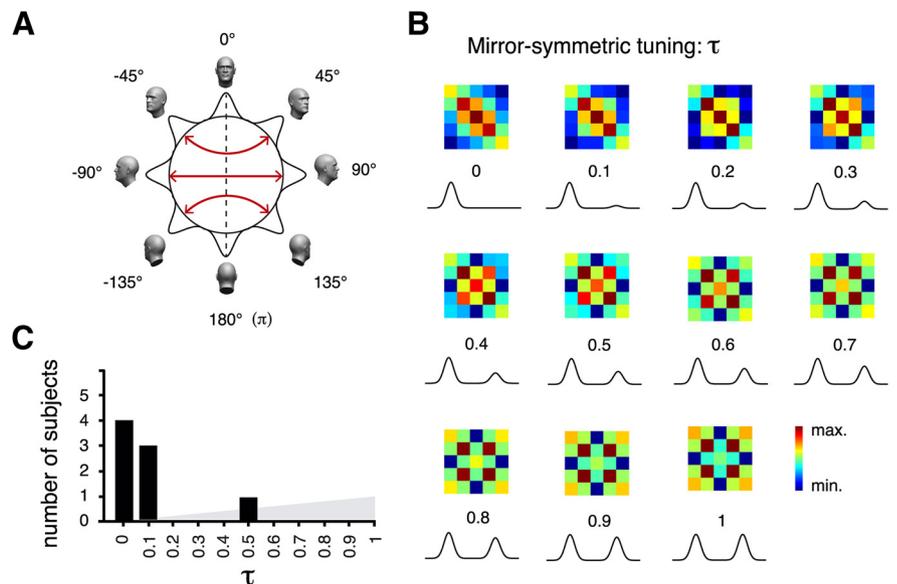
We modeled multivariate responses in rFFA as feature vectors consisting of as many elements as the mean number of voxels in our data. The response level in each voxel to our experimental conditions was modeled as the total activation of several neural populations, each preferentially responding to a specific view with a Gaussian tuning curve, the total range of views spanning a full rotation of the head (Fig. 6*A,B*). To capture effects of SNR on the parametric family of mSMs, i.i.d. Gaussian noise was added to the simulated response patterns. The space of the model parameters (population tuning width ( $\sigma$ ), degree of front-view preference ( $K$ ), and level of clustering ( $n_{patch}$ ) of similarly view-tuned neurons assumed to reside within each FFA voxel) was exhaustively covered to generate a set of mSMs. The statistical expectation of mSMs for several combinations of  $\sigma$  and  $K$  parameter values are shown in Figure 6*B*. Then, we found the best fitting “model template” for each subject, defined as that mSM displaying the largest correlation with a subject’s eSM. Best fitting  $\sigma$  and  $K$  parameter values for each subject are plotted in Figure 6*C*. The observed clustering of points in the lower-right quadrant indicates: (1) An overrepresentation of the frontal views of faces (sign-permutation test,  $p = 0.0156$ ), and (2) a population tuning width of  $\sim 30^\circ$  (range,  $20^\circ, 40^\circ$ ).

**Model extension: probing for mirror-symmetric tuning**

Next, we explored a variant of our model incorporating mirror symmetry directly as a parameter (Fig. 7). From a descriptive point of view, the model produced distinguishable mSMs for differing degrees of mirror-symmetric tuning, including mSMs strongly correlated to those used for RSA. Nevertheless, mirror-symmetric tuning as estimated by the model was found to be low (mean 10%). Because, to our knowledge, it is not possible to test the null hypothesis of  $\tau = 0$  (i.e., no mirror symmetry) because the parameter is bounded at 0 and, further, because failing to reject a null hypothesis cannot demonstrate the absence of an effect, we rather adopted a model comparison approach. From a model comparison perspective, our data in rFFA were not better accounted for by the model including mirror-symmetric tuning compared with the simpler model assuming unimodal tuning functions. Cross-validated model comparison resulted in only 3 of 8 votes in favor of the



**Figure 6.** Computational modeling. *A*, fMRI responses in FFA were modeled based on the inhomogeneous sampling by fMRI voxels of view-tuned clusters of cells covering a full rotation of the head. *B*, Model parameters included tuning width ( $\sigma$ ), view-tuning distribution ( $K$ ), and patchiness (size) of neural clusters ( $n_{patch}$ ). *C*, Expected dissimilarity matrices for different combinations of parameters  $\sigma$  and  $K$ . *D*, Model fits for eight subjects revealing an overrepresentation of the front view (i.e.,  $K > 0$ ).



**Figure 7.** Computational model including mirror-symmetric tuning. *A*, fMRI responses in FFA were simulated based on the inhomogeneous sampling by fMRI voxels of cell clusters with one of many possible degrees of mirror-symmetric tuning. Mirror-symmetric views are indicated by red arrows. The dotted-line indicates the axis of symmetry. *B*, Expected model similarity matrices for increasing degrees of mirror-symmetric tuning as defined by parameter  $\tau$ . Distinct similarity structures associated to increasing levels of  $\tau$  can be observed. Tuning functions below each similarity matrix represent the associated degree of mirror-symmetric tuning. *C*, Histogram of the fitted  $\tau$  parameters of eight subjects revealing generally low levels of mirror symmetry. Mean  $\tau$  across the population = 0.1.

model including mirror-symmetric tuning. In sum, we did not observe in rFFA conclusive evidence of mirror symmetry. Model comparison did not favor as a characterization of our data the more complex model, including various degrees of mirror-symmetric tuning over the simpler model assuming unimodal tuning functions.

## Discussion

A combination of fMRI and MVPA was used to show that, although LO implements a category-general shape-code that partially generalizes across retinal positions, rFFA implements a face-selective and translation-tolerant representation of orientation. Critically, encoding in FFA was consistent with a monotonic code reflecting head angular disparity, and not with a mirror-symmetric scheme when using an analysis that is not sensitive to mean activation differences between conditions. We relied on computational modeling to demonstrate that inhomogeneously distributed clusters of similarly view-tuned neurons, approximately analogous to those observed in monkey inferotemporal cortex (Fujita et al., 1992; Wang et al., 1996), can explain the representational structure of orientation information observed in FFA. Modeling results also revealed a stronger representation of the frontal views of faces.

Our classification results agree with previous fMRI adaptation (fMRIa) studies of ventral visual cortex showing a release from adaptation for faces and objects rotated in-depth (Grill-Spector et al., 1999; Gauthier et al., 2002; Andrews and Ewbank, 2004; Fang et al., 2007b; Andresen et al., 2009; Kim et al., 2009) and confirm recent MVPA studies reporting face orientation information in FFA (Axelrod and Yovel, 2012; Kietzmann et al., 2012). However, our findings go beyond previous studies in two fundamental ways. First, because we obtained data for two distinct object categories, we were able to investigate the category specificity of orientation information. Thus, in contradistinction to LO, where orientation was decoded with comparable accuracy for faces and vehicles, in rFFA we found face-selective encoding of orientation information. Second, because we presented stimuli in two distinct retinal locations, we were able to probe the generalizability of these representations across retinal positions. Thus, we obtained novel evidence that in rFFA head orientation information largely generalizes across positions. Previous fMRIa studies have provided mixed evidence concerning the category specificity of orientation encoding in FFA (Andrews and Ewbank, 2004; Fang et al., 2007b; Andresen et al., 2009).

A difference between previous MVPA studies (Axelrod and Yovel, 2012; Kietzmann et al., 2012) and our findings concerns the representational structure of orientation information in FFA. Whereas our data support a monotonic code presumably reflecting neurons unimodally tuned to a single preferred view, the other two studies also suggest some evidence of additional effects of mirror symmetry, putatively reflecting bimodally tuned cells.

Our main analysis aimed to avoid effects of the overall mean in the FFA on rotational angle. The reason is that the mean effect is sensitive to many parameters (in addition to rotation). In particular, it distinguishes between faces and nonfaces and is also sensitive to the size and contrast of faces (e.g., Yue et al., 2011). Therefore, if FFA encoded face rotational angle or face identity using the mean, it would not be possible to distinguish face orientation or identity from changes in contrast or size. It is a common criterion in visual neuroscience that object representations are invariant or at least tolerant to differences in low-level features (Sáry et al., 1993; Tovee et al., 1994; Ito et al., 1995; Freiwald and Tsao, 2010), which does not hold for the mean signal. Furthermore, the tuning and read-out of information encoded in voxel ensembles only indirectly points toward coding at the level of populations of single neurons.

However, when using a different, Euclidian distance-based RSA approach that is more sensitive to effects of the mean, we find that the mirror-symmetric and the monotonic model re-

flecting angular distance are indistinguishable. Thus, to some degree, the differences between our study and previous work (Axelrod and Yovel, 2012; Kietzmann et al., 2012) can depend on analysis choices. These choices may also reflect differences in theoretical concepts regarding neural representation (e.g., whether to include the mean as a coding dimension, as discussed above). A full comparison of all the effects of similarities and differences in analysis choices between our paper and previous studies is beyond the scope of a single empirical study.

A critical assumption behind studies combining neuroimaging and MVPA is that signatures of neural tuning functions are discoverable on the basis of large-scale brain activity measures. However, it is discriminative information that MVPA detects, not tuning properties of neurons. To narrow this gap, we investigated whether the monotonic model favored by RSA can indeed be recovered from fMRI signals under biologically plausible assumptions. Thus, using a novel modeling approach, we show that unimodally tuned neural responses to face views, as predominantly observed in macaque temporal cortex (Perrett et al., 1991, 1998; Logothetis et al., 1995), and a columnar organization approximately similar to that observed in nonhuman primates (Fujita et al., 1992; Wang et al., 1996), is consistent with our observations. Thus, from distributed patterns alone, we detected a more robust representation of the front views of faces, in line with previous univariate fMRI (Andresen et al., 2009; Yue et al., 2011; Axelrod and Yovel, 2012) and nonhuman primate electrophysiological studies (Hasselmo et al., 1989; Perrett et al., 1991, 1998). However, before strong claims are made regarding an effective prevalence of neurons tuned to frontally viewed faces in FFA, further research is needed concerning the effect of face orientation on FFA localization. Given that our independent localizer run consisted of frontally viewed faces (standard practice in the field), we cannot discard that the detected overrepresentation may partially reflect this bias. Nevertheless, regardless of its underlying cause, our results show that, in rFFA, an overrepresentation of frontally viewed faces is reflected in the corresponding fMRI patterns. Further, this suggests lower SNRs for nonfrontally viewed faces, a fact that may have influenced findings of mirror symmetry.

The population tuning-width  $\sigma$  estimated by our method rests on assumptions about neural clustering that are not yet verified in humans. Furthermore, our method cannot be treated as a quantitative measure and therefore should be interpreted carefully. Nonetheless, the estimated population tuning width of 35° (HWHM) is approximately consistent with electrophysiological recordings; the majority of face-selective neurons in macaque temporal cortex are unimodally tuned to a single preferred view (>80%), the majority of these cells exhibiting tuning-widths <60° (Perrett et al., 1991). However, reported tuning widths to faces and objects in monkey temporal lobe range from narrow (20°–40°) to broad (~100°) (Perrett et al., 1991; Logothetis et al., 1995; Eifuku et al., 2004). Intriguingly, although Freiwald and Tsao (2010) observed a broad range of angular tuning widths for single neurons in MFP, our results are approximately consistent with their observations at the population level. Our estimated population tuning-width is also in agreement with channel bandwidths suggested by psychophysics; view-aftereffects abruptly decrease beyond disparities of 30° (Fang and He, 2005; Jeffery et al., 2007). Interestingly, evidence suggests these view-aftereffects do not generalize across categories (faces and vehicles), are tolerant to retinal translation, and are reduced for inverted faces (Fang and He, 2005; Fang et al., 2007a; Jeffery et al., 2007).

In sum, our modeling approach provides a biologically plausible explanation of our findings, accounts for the effect of unequal SNRs on empirical dissimilarity matrices, and makes predictions concerning neural tuning properties that can be directly tested by other neuroscientific methods. Thus, our approach might afford a rich tool for the interpretation of population codes detected with MVPA and help investigate the underlying neural populations. Further, our findings suggest that the detectability of information by fMRI combined with MVPA depends on the spatial relationships among neurons as well as their tuning functions, perhaps precluding the detection of sparsely coded information.

Although homologies of visual areas beyond V4 are debated (Tsao et al., 2008; Pinsk et al., 2009), our findings agree with the hypothesized homology between the macaque MFP and FFA of Tsao et al. (2003). Crucially, Freiwald and Tsao (2010) found unimodally tuned responses in MFP, in agreement with our observations in FFA. If further homologies hold, our findings would imply a strongly mirror-symmetrically tuned processing stage in humans located anterior to FFA and posterior to aIT, the latter shown to encode face identity (Kriegeskorte et al., 2007). This putative face-sensitive area may correspond to the region observed by Tsao et al. (2008) and Rajimehr et al. (2009), and also approximately coincides with an area slightly anterior to FFA found to encode face identity information (Nestor et al., 2011). However, given that only frontal faces were tested in these studies (Tsao et al., 2008; Rajimehr et al., 2009; Nestor et al., 2011), the crucial prediction of increased similarities for mirror-symmetric face views in this region compared with FFA remains to be tested.

In conclusion, previous studies suggest that invariant object recognition relies on a hierarchical architecture along the ventral stream. Consistent with this, we show that translation-tolerant shape information, independent of object category, is available in LO. Moreover, our findings suggest that rFFA implements a category-selective and translation-invariant representation of orientation. The structure of this representation proved consistent with a monotonic angle-code, especially when the mean signal in rFFA was excluded as a coding dimension. This may reflect unimodally tuned clusters of neurons, in line with observations in monkey IT.

## References

- Andresen DR, Vinberg J, Grill-Spector K (2009) The representation of object viewpoint in human visual cortex. *Neuroimage* 45:522–536. [CrossRef Medline](#)
- Andrews TJ, Ewbank MP (2004) Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage* 23:905–913. [CrossRef Medline](#)
- Axelrod V, Yovel G (2012) Hierarchical processing of face viewpoint in human visual cortex. *J Neurosci* 32:2442–2452. [CrossRef Medline](#)
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Bülthoff HH, Edelman S (1992) Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proc Natl Acad Sci U S A* 89:60–64. [CrossRef Medline](#)
- Bülthoff HH, Edelman SY, Tarr MJ (1995) How are three-dimensional objects represented in the brain? *Cereb Cortex* 5:247–260. [CrossRef Medline](#)
- Carlson TA, Schrater P, He S (2003) Patterns of activity in the categorical representations of objects. *J Cogn Neurosci* 15:704–717. [CrossRef Medline](#)
- Carlson T, Hogendoorn H, Fonteijn H, Verstraten FA (2011) Spatial coding and invariance in object-selective cortex. *Cortex* 47:14–22. [CrossRef Medline](#)
- Cichy RM, Chen Y, Haynes JD (2011) Encoding the identity and location of objects in human LOC. *Neuroimage* 54:2297–2307. [CrossRef Medline](#)
- Cichy RM, Sterzer P, Heinzle J, Elliott LT, Ramirez F, Haynes JD (2013) Probing principles of large-scale object representation: category preference and location encoding. *Hum Brain Mapp* 34:1636–1651. [CrossRef Medline](#)
- Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage* 19:261–270. [CrossRef Medline](#)
- DiCarlo JJ, Zoccolan D, Rust NC (2012) How does the brain solve visual object recognition? *Neuron* 73:415–434. [CrossRef Medline](#)
- Efron B, Tibshirani RJ (1994) *An introduction to the bootstrap*, Ed 1. Boca Raton, FL: Chapman and Hall/CRC.
- Eger E, Ashburner J, Haynes JD, Dolan RJ, Rees G (2008a) fMRI activity patterns in human LOC carry information about object exemplars within category. *J Cogn Neurosci* 20:356–370. [CrossRef Medline](#)
- Eger E, Kell CA, Kleinschmidt A (2008b) Graded size sensitivity of object-exemplar-evoked activity patterns within human LOC subregions. *J Neurophysiol* 100:2038–2047. [CrossRef Medline](#)
- Eifuku S, De Souza WC, Tamura R, Nishijo H, Ono T (2004) Neuronal correlates of face identification in the monkey anterior temporal cortical areas. *J Neurophysiol* 91:358–371. [CrossRef Medline](#)
- Fang F, He S (2005) Viewer-centered object representation in the human visual system revealed by viewpoint aftereffects. *Neuron* 45:793–800. [CrossRef Medline](#)
- Fang F, Ijichi K, He S (2007a) Transfer of the face viewpoint aftereffect from adaptation to different and inverted faces. *J Vis* 7:6. [CrossRef Medline](#)
- Fang F, Murray SO, He S (2007b) Duration-dependent fMRI adaptation and distributed viewer-centered face representation in human visual cortex. *Cereb Cortex* 17:1402–1411. [CrossRef Medline](#)
- Freiwald WA, Tsao DY (2010) Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* 330:845–851. [CrossRef Medline](#)
- Fujita I, Tanaka K, Ito M, Cheng K (1992) Columns for visual features of objects in monkey inferotemporal cortex. *Nature* 360:343–346. [CrossRef Medline](#)
- Gauthier I, Hayward WG, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (2002) BOLD activity during mental rotation and viewpoint-dependent object recognition. *Neuron* 34:161–171. [CrossRef Medline](#)
- Good PI (2004) *Permutation, parametric, and bootstrap tests of hypotheses*, Ed 3. New York: Springer.
- Grill-Spector K, Malach R (2004) The human visual cortex. *Annu Rev Neurosci* 27:649–677. [CrossRef Medline](#)
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24:187–203. [CrossRef Medline](#)
- Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. *Vision Res* 41:1409–1422. [CrossRef Medline](#)
- Grill-Spector K, Knouf N, Kanwisher N (2004) The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci* 7:555–562. [CrossRef Medline](#)
- Hansen KA, David SV, Gallant JL (2004) Parametric reverse correlation reveals spatial linearity of retinotopic human V1 BOLD response. *Neuroimage* 23:233–241. [CrossRef Medline](#)
- 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. [Medline](#)
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430. [CrossRef Medline](#)
- Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8:686–691. [CrossRef Medline](#)
- Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. *Nat Rev Neurosci* 7:523–534. [CrossRef Medline](#)
- Ito M, Tamura H, Fujita I, Tanaka K (1995) Size and position invariance of neuronal responses in monkey inferotemporal cortex. *J Neurophysiol* 73:218–226. [Medline](#)
- Jeffery L, Rhodes G, Busey T (2007) Broadly tuned, view-specific coding of face shape: opposing figural aftereffects can be induced in different views. *Vision Res* 47:3070–3077. [CrossRef Medline](#)
- Jeffreys H (1961) *Theory of probability*, Ed 3. Oxford: Oxford UP.

- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685. [CrossRef Medline](#)
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311. [Medline](#)
- Kietzmann TC, Swisher JD, König P, Tong F (2012) Prevalence of selectivity for mirror-symmetric views of faces in the ventral and dorsal visual pathways. *J Neurosci* 32:11763–11772. [CrossRef Medline](#)
- Kim JG, Biederman I, Lescroart MD, Hayworth KJ (2009) Adaptation to objects in the lateral occipital complex (LOC): shape or semantics? *Vision Res* 49:2297–2305. [CrossRef Medline](#)
- Kriegeskorte N (2009) Relating population-code representations between man, monkey, and computational models. *Front Neurosci* 3:363–373. [CrossRef Medline](#)
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868. [CrossRef Medline](#)
- Kriegeskorte N, Formisano E, Sorger B, Goebel R (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc Natl Acad Sci U S A* 104:20600–20605. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis: connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:1–28. [CrossRef Medline](#)
- Larsson J, Heeger DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci* 26:13128–13142. [CrossRef Medline](#)
- Logothetis NK, Pauls J, Poggio T (1995) Shape representation in the inferior temporal cortex of monkeys. *Curr Biol* 5:552–563. [CrossRef Medline](#)
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A* 92:8135–8139. [CrossRef Medline](#)
- Nestor A, Plaut DC, Behrmann M (2011) Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proc Natl Acad Sci U S A* 108:9998–10003. [CrossRef Medline](#)
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430. [CrossRef Medline](#)
- Perrett DI, Oram MW, Harries MH, Bevan R, Hietanen JK, Benson PJ, Thomas S (1991) Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Exp Brain Res* 86:159–173. [Medline](#)
- Perrett DI, Oram MW, Ashbridge E (1998) Evidence accumulation in cell populations responsive to faces: an account of generalisation of recognition without mental transformations. *Cognition* 67:111–145. [CrossRef Medline](#)
- Pinsk MA, Arcaro M, Weiner KS, Kalkus JF, Inati SJ, Gross CG, Kastner S (2009) Neural representations of faces and body parts in macaque and human cortex: a comparative fMRI study. *J Neurophysiol* 101:2581–2600. [CrossRef Medline](#)
- Rajimehr R, Young JC, Tootell RB (2009) An anterior temporal face patch in human cortex, predicted by macaque maps. *Proc Natl Acad Sci U S A* 106:1995–2000. [CrossRef Medline](#)
- Sáry G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260:995–997. [CrossRef Medline](#)
- Sayres R, Grill-Spector K (2008) Relating retinotopic and object-selective responses in human lateral occipital cortex. *J Neurophysiol* 100:249–267. [CrossRef Medline](#)
- Schwarzlose RF, Swisher JD, Dang S, Kanwisher N (2008) The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci U S A* 105:4447–4452. [CrossRef Medline](#)
- Tarr MJ, Williams P, Hayward WG, Gauthier I (1998) Three-dimensional object recognition is viewpoint dependent. *Nat Neurosci* 1:275–277. [CrossRef Medline](#)
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381:520–522. [CrossRef Medline](#)
- Tong F, Nakayama K, Moscovitch M, Weinrib O, Kanwisher N (2000) Response properties of the human fusiform face area. *Cogn Neuropsychol* 17:257–280. [CrossRef Medline](#)
- Tovee MJ, Rolls ET, Azzopardi P (1994) Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. *J Neurophysiol* 72:1049–1060. [Medline](#)
- Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB (2003) Faces and objects in macaque cerebral cortex. *Nat Neurosci* 6:989–995. [CrossRef Medline](#)
- Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. *Science* 311:670–674. [CrossRef Medline](#)
- Tsao DY, Moeller S, Freiwald WA (2008) Comparing face patch systems in macaques and humans. *Proc Natl Acad Sci U S A* 105:19514–19519. [CrossRef Medline](#)
- Ullman S (1998) Three-dimensional object recognition based on the combination of views. *Cognition* 67:21–44. [CrossRef Medline](#)
- Vapnik VN (1995) *The nature of statistical learning theory*. New York: Springer.
- Wang G, Tanaka K, Tanifuji M (1996) Optical imaging of functional organization in the monkey inferotemporal cortex. *Science* 272:1665–1668. [CrossRef Medline](#)
- Wang G, Tanifuji M, Tanaka K (1998) Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. *Neurosci Res* 32:33–46. [CrossRef Medline](#)
- Yue X, Cassidy BS, Devaney KJ, Holt DJ, Tootell RB (2011) Lower-level stimulus features strongly influence responses in the fusiform face area. *Cereb Cortex* 21:35–47. [CrossRef Medline](#)