

# Cortical Thickness in Fusiform Face Area Predicts Face and Object Recognition Performance

Rankin W. McGugin<sup>1</sup>, Ana E. Van Gulick<sup>2</sup>, and Isabel Gauthier<sup>1</sup>

## Abstract

■ The fusiform face area (FFA) is defined by its selectivity for faces. Several studies have shown that the response of FFA to nonface objects can predict behavioral performance for these objects. However, one possible account is that experts pay more attention to objects in their domain of expertise, driving signals up. Here, we show an effect of expertise with nonface objects in FFA that cannot be explained by differential attention to objects of expertise. We explore the relationship between cortical thickness of FFA and face and object recognition using the Cambridge Face Memory Test and Vanderbilt Expertise Test, respectively. We

measured cortical thickness in functionally defined regions in a group of men who evidenced functional expertise effects for cars in FFA. Performance with faces and objects together accounted for approximately 40% of the variance in cortical thickness of several FFA patches. Whereas participants with a thicker FFA cortex performed better with vehicles, those with a thinner FFA cortex performed better with faces and living objects. The results point to a domain-general role of FFA in object perception and reveal an interesting double dissociation that does not contrast faces and objects but rather living and nonliving objects. ■

## INTRODUCTION

Functional brain imaging research has offered strong support for localized functions in the brain. However, brain imaging findings often generate debate with respect to the attribution of specific cognitive functions to patterns of localized responses (Grodzinsky & Santi, 2008; Shomstein & Yantis, 2006; Price & Devlin, 2003; Burton, Small, & Blumstein, 2000). For instance, should we conceive of the FFA as a specialized module dedicated only to the processing of faces, with little, if any, role in the processing of other objects (Kanwisher, 2010)? Or can we understand the strong selectivity for faces in FFA as resulting from expertise with faces, such that other objects with similar experience would also recruit the FFA (Tarr & Gauthier, 2000)? Questioning the evidence of domain specificity in FFA is questioning some of the strongest evidence of domain specificity in the visual system and the brain.

Fifteen years past the first experiment reporting expertise effects in FFA after training with novel objects called Greebles (Gauthier & Tarr, 1997), several studies of individual variability in FFA BOLD response in real-world domains suggest that the response of FFA to nonface objects can predict behavioral performance for these objects (e.g., McGugin, Newton, Gore, & Gauthier, 2014; Bilalić, Langner, Ulrich, & Grodd, 2011; Xu, 2005; Gauthier, Skudlarski, Gore, & Anderson, 2000). Expertise effects

are obtained in the very middle of the FFA (McGugin, Van Gulick, Tamber-Rosenau, Ross, & Gauthier, 2014), even in the most highly face-selective voxels in high-resolution (HR) scans (McGugin, Gatenby, Gore, & Gauthier, 2012). However, other studies have found no correlation between performance with cars and FFA response (e.g., Grill-Spector, Knouf, & Kanwisher, 2004) or failed to replicate the Greeble training effect (Brants, Wagemans, & Op de Beeck, 2011).<sup>1</sup>

One concern about expertise effects in the visual system is that they may be because of greater attention to objects of expertise (Harel, Gilaie-Dotan, Malach, & Bentin, 2010). This account has been challenged by demonstrations of robust expertise effects in FFA under conditions that reduce these effects in other visual areas (McGugin, Newton, et al., 2014; McGugin, Van Gulick et al., 2014). However, attention is a strong modulator of responses in visual cortex (Pessoa, Kastner, & Ungerleider, 2003), and it is plausible for people to pay more attention to objects of expertise (including faces). An attentional account of expertise effects of functional MRI data is difficult to rule out entirely.

Here, we turn to the study of the structural correlates of face and object recognition ability and note that such expertise effects, whether they are related to functional effects, could not be explained by attention. Test-retest reliability of structural MRI data and, specifically, surface maps of CT are highly reproducible with high intraclass correlations (Wonderlick et al., 2009), allowing us to comfortably look at individual differences in regional CT. Measures of regional brain structure have been successfully

<sup>1</sup>Vanderbilt University, <sup>2</sup>Carnegie Mellon University

associated with performance in a number of domains (Delon-Martin, Plailly, Fonlupt, Veyrac, & Royet, 2013; Schwarzkopf, Song, & Rees, 2011; Foster & Zatorre, 2010; Karama et al., 2009; Wong et al., 2008; Narr et al., 2007; Hyde, Zatorre, Griffiths, Lerch, & Peretz, 2006; Shaw et al., 2006; Schneider et al., 2005; Golestani, Paus, & Zatorre, 2002). These studies demonstrate individual differences in brain structure in the same areas where differences in BOLD activation are seen, and both types of brain reorganization are associated with domain-specific behavioral differences. Accordingly, we may expect CT in FFA to be related to behavioral face recognition performance (McGugin, Gatenby, et al., 2012; Xu, 2005; Grill-Spector et al., 2004; Kanwisher, McDermott, & Chun, 1997).

In one study with prosopagnosic patients, the right fusiform gyrus showed reduced gray matter volume relative to normal controls (Garrido et al., 2009). However, using healthy participants, recent work (Bi, Chen, Zhou, He, & Fang, 2014) found a negative correlation between cortical thickness (CT) in left FFA (lFFA) and improvements in a task involving judging the orientation of faces. This was not a face recognition task, and so it is unclear whether face performance should also show the same negative correlation with CT or follow the general trend observed when performance in patients versus controls is correlated with BOLD response.

We might also expect CT in FFA to be related to object recognition performance, based on functional effects of expertise in this region. However, one report found that expertise with cars was related to gray matter volume in pFC but not in the fusiform gyrus (Gilaie-Dotan, Harel, Bentin, Kanai, & Rees, 2012). We chose to revisit this question because the aforementioned study used a group-averaged template, as is typical in brain morphometry, to look for brain areas whose structure might be related to behavior. Even when functional ROIs have been used in studies looking at brain structure (Bi et al., 2014), they have typically been group-averaged ROIs. Within the fusiform gyrus, functional effects of expertise are spatially limited to two small face-selective areas (Weiner et al., 2014) and are best revealed in individually defined ROIs.

We performed CT analyses in individually defined functional ROIs in a sample of 27 men who were recruited to vary in their expertise for cars. We defined ROIs functionally and individually. None of the prior work with CT used individual functional ROIs. In addition, our structural scans come from a sample of participants who showed the expected positive correlation between behavioral performance with cars and FFA selectivity to cars in a prior study (McGugin, Van Gulick, et al., 2014). Therefore, we are able to ask if CT predicts behavioral performance in participants whose performance with cars was related to the BOLD selectivity for cars. Critically, however, there is no reason why CT should be specifically related to the object category(ies) used in our separate functional task. Brain structure could be related to per-

formance with any object category. For this reason, we used behavioral performance for a variety of object categories and faces, in a battery of visual learning tasks (the Vanderbilt Expertise Test, VET; McGugin, Richler, Herzmann, Speegle, & Gauthier, 2012) and the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006). VET performance for vehicles shows a stronger relationship with the CFMT in men than women (McGugin, Richler, et al., 2012). Because of such gender differences and because the sample we used was composed of men (gender has too large of an effect on CT to justify including the three women in the original McGugin, Gatenby, et al. [2012] study), we decided to index object recognition performance according to the two principle factors extracted from a principal component analysis of the VET results, which, in prior work, also correlated with gender. The first factor corresponds to living objects (on which women generally performed better than men), and the second corresponds to nonliving objects (on which men generally performed better than women; McGugin, Richler, et al., 2012). Thus, the behavioral indices of performance used here are the same measures as in several studies of expertise (McGugin, Newton, et al., 2014; McGugin, Van Gulick, et al., 2014; Curby, Glazek, & Gauthier, 2009; Gauthier, Curby, Skudlarski, & Epstein, 2005; Xu, 2005; Grill-Spector et al., 2004; Rossion, Kung, & Tarr, 2004; Gauthier et al., 2000). We average categories for which performance tends to be correlated, which may help detect small effects associated with each category. Because this study sample was recruited with regard to their car expertise, we also investigate correlations with car performance alone.

We hypothesized that we would find linear relationships between CT in FFA and performance for both faces and objects. Importantly, the literature contains examples of better performance in various domains that are associated with either thicker (Foster & Zatorre, 2010; Karama et al., 2009; Choi et al., 2008; Narr et al., 2007) or thinner (Jung et al., 2010; Hyde et al., 2007) cortex. For this reason, we do not formulate a prediction for the direction of the linear relations between performance and local CT, and we use two-tailed tests.

## METHODS

### Participants

Twenty-seven healthy right-handed men (age: range = 18–34 years, mean =  $26 \pm 4.7$  years) participated as volunteers for a larger study that also included three women, aimed at investigating effects of behavioral expertise under conditions of visual clutter (McGugin, Van Gulick, et al., 2014). The current work represents a new analysis of the structural data that was used in McGugin, Van Gulick, et al. (2014) only as support for functional analyses. Informed written consent was obtained from each participant in accordance with guidelines of

the institutional review board of Vanderbilt University and Vanderbilt University Medical Center. All participants received monetary compensation for their participation and had normal or corrected-to-normal vision. One participant was discarded because of outlier performance (at or below chance of .33) for six of the eight object categories in the behavioral memory test.

### Behavioral Assessments

All participants completed three behavioral tasks outside the scanner: the CFMT (Duchaine & Nakayama, 2006), the VET (McGugin, Richler, et al., 2012), and a sequential matching expertise test used to quantify individual skill at matching cars (McGugin, Gatenby, et al., 2012; Curby et al., 2009; Gauthier et al., 2000, 2005; Xu, 2005; Grill-Spector et al., 2004; Rossion et al., 2004). See Table 1 for descriptive statistics for all behavioral measures.

In the CFMT, participants study three images (left one-third profile, frontal view, and right one-third profile) of the first target face for 3 sec per image, immediately followed by three test items where participants select the studied image among two distractors. This introductory learning phase is repeated for the remaining five target faces. Participants were then presented with 30 forced-choice test displays, each containing one target face and two distractor faces. Participants were instructed to select the face that matched one of the original six target faces. The matching faces varied from their original presentation by means of lighting, pose, or both. Next, participants were again presented with the six target faces to study, followed by 24 test displays presented in Gaussian noise. For a complete description of the CFMT, see Duchaine and Nakayama (2006).

The VET (McGugin, Richler, et al., 2012) includes eight object categories blocked alphabetically: butterflies, cars, leaves, motorcycles, mushrooms, owls, planes, and wading birds. For each category, participants studied a display with images from each of six species/models. For each test trial, one of the studied targets (identical images for the first 12 trials or transfer images requiring generalization across viewpoint, size, and settings for the subsequent 36 trials) was presented with two distractors from another species/model in a forced-choice paradigm. The target image could occur in any of the three positions, and participants indicated which image of the triplet was the studied target. Before beginning the VET, participants rated themselves on their expertise with all tested categories (leaves, owls, butterflies, wading birds, mushrooms, cars, planes, and motorcycles) and also with faces, considering “interest in, years exposure to, knowledge of, and familiarity with each category,” where 1 represented the lowest reported skill level and 9 represented the highest. See Table 1 for descriptive statistics of self-report (SR) scores. For a complete description of the VET, see McGugin, Richler, et al. (2012).

**Table 1.** Descriptive Statistics

	<i>Average CT (SD)</i>	<i>Range</i>
rFFA1	2.4 (0.3)	1.53–2.95
rFFA2	2.7 (0.3)	2.17–3.58
Right OFA	2.4 (0.7)	1.75–5.05
Right PHG	2.4 (0.4)	1.38–3.25
lFFA1	2.3 (0.4)	1.43–3.71
lFFA2	2.6 (0.5)	1.44–3.51
Left OFA	2.2 (0.4)	1.21–3.1
Left PHG	2.5 (0.4)	1.8–3.94
Global CT	2.5 (0.3)	1.8–2.9
Age	26 (4.7)	18–34
CFMT_all	0.8 (0.1)	0.57–0.96
VET-LV	0.6 (0.1)	0.47–0.69
VET-NL	0.7 (0.1)	0.40–0.84
Matching bird	1.3 (0.4)	0.56–2.14
Matching car/plane	0.6 (0.7)	–0.54 to 2.17
SR-LV	2.2 (1)	1–5.2
SR-NL	4.2 (1.7)	1–7.67

Principal component analysis has demonstrated that the underlying structure of the eight-category VET is largely explained by two independent factors that represent living and nonliving objects. Therefore, we reduced VET performance to a living objects score (VET-LV; average of butterflies, leaves, mushrooms, owls, and wading birds) and a nonliving objects score (VET-NL; average of cars, motorcycles, and planes).

The matching task has 112 sequential matching trials for each of the three categories: cars, planes, and birds (56 unique images/category). On each trial, a first stimulus appeared for 1000 msec, followed by a 500-msec mask and second stimulus that remained visible until participants made a same or different response or 5000 msec elapsed. Participants judged if the two images showed cars/planes of the same make and model regardless of year or birds of the same species.

### MRI Acquisition

Scanning was performed using a Philips (Amsterdam, The Netherlands) 3-T Intera Achieva MRI scanner with an eight-channel head coil located at the Vanderbilt University Institute for Imaging Science. HR T1-weighted anatomical volumes were acquired (repetition time = 8.93 msec; echo time = 4.6 msec; flip angle = 9°; field of view [FOV] = 256 × 256; slice thickness = 1 mm, no gap; in-plane resolution = 1 × 1 mm; 170 slices acquired in the sagittal plane). In a functional localizer run, we used standard gradient-echo

echo planar T2\*-weighted imaging to obtain functional images (repetition time = 2000 msec; echo time = 35 msec; flip angle = 79°; FOV = 192 × 192; slice thickness = 3 mm, no gap; in-plane resolution = 3 × 3 mm; 34 ascending interleaved slices acquired axially).

The structural scan was processed using Brain Voyager v2.6 (Maastricht, The Netherlands, [www.brainvoyager.com](http://www.brainvoyager.com)). First, steps were taken to prepare the brain for automatic correction of intensity inhomogeneities; the image background was cleaned, the brain was extracted, and the bias field was estimated and removed. The cerebellum and brainstem were manually removed for each brain. After automatic intensity inhomogeneity correction, the gray matter and white matter intensities were centered around intensity values of 100 and 160, respectively. Brains were then Talairach normalized and interpolated to 0.5 × 0.5 × 0.5 mm resolution. The white/gray matter boundary was segmented, after which the gray matter/cerebrospinal fluid boundary (corresponding to the pial surface or the outer boundary of the cortex) was labeled.

For the functional localizer scan, all images were presented with an Apple Macintosh computer running MATLAB (The MathWorks, Natick, MA) using the Psychophysics Toolbox extension (Santa Barbara, CA, Brainard, 1997; Pelli, 1997). Stimuli were displayed on a rear-projection screen using an Eiki LC-X60 LDP projector with a Navitar zoom lens (Rancho Santa Margarita, CA). Seventy-two grayscale images (36 faces, 36 objects) were used in a 1-back detection task with 18 alternating blocks of faces or objects (16 images shown for 1 sec) and a 2-sec fixation at the beginning and end of each block. Sensitivity did not differ for face and object blocks: hit rate and false alarm rate, face = 0.92 and 0.008 and object = 0.93 and 0.004.

After the functional localizer scan, participants completed eight runs using different combinations of images and tasks (see McGugin, Van Gulick, et al., 2014, for full details). To verify the face selectivity of the ROIs in this subset of participants, we analyzed only the first two of these experimental runs to obtain an independent mea-

sure of face selectivity in the ROIs defined in the functional localizers. These runs showed single objects presented in isolation in a blocked fMRI design with a 1-back repetition task of face, car, or butterfly images.

## Data Analysis

The HR T1-weighted structural scans were normalized to Talairach space. Functional data were analyzed using Brain Voyager ([www.brainvoyager.com](http://www.brainvoyager.com)) and in-house MATLAB scripts. Preprocessing included registration to the original (nontransformed) structural scan, slice scan time correction (cubic spline), 3-D motion correction (trilinear/sinc interpolation), and temporal filtering (high-pass criterion of two cycles per run) with linear trend removal.

ROIs were defined using the face > object contrast from the face localizer scan (Table 2). For ROI analyses, no spatial smoothing was applied to the CT maps. We localized bilateral ROIs that responded more to faces than objects in the posterior fusiform gyrus (FFA1), middle fusiform gyrus (FFA2; Weiner, Sayres, Vinberg, & Grill-Spector, 2010; Pinsk et al., 2009), and occipital face area (OFA) and more to objects than faces in the parahippocampal gyrus (PHG). To verify the face selectivity of these regions using functional data independent from the localizer, we examined the BOLD response to faces relative to a butterfly baseline (cars were not used because several participants were car experts). As expected, there was a larger response to faces versus butterflies in bilateral FFA1, FFA2, and OFA and the opposite effect in object-defined regions in the PHG (Table 2).

All ROIs were initially defined based on the 1-mm (interpolated) statistical maps using a fixed millimeter spread of activation to ensure consistency with reported sizes of these functional ROIs in the literature as well as consistency across participants (Table 1). However, to ensure that the signal was weighted per functional voxel, ROIs were subsequently downsampled to functional (3-mm) resolution. Any functional voxel containing one or more 1-mm voxel from the initial ROI was considered to be part of the final

**Table 2.** Localization of ROIs

	<i>N</i>	<i>Mean Talairach Coordinates for Peak Face-selective Voxel ± SD</i>	<i>Face Selectivity (Face–Butterfly; 95% CI)</i>	<i>t Test of Mean Face Selectivity: t Statistic (p Value)</i>
rFFA1	26	40, –59, –23 (4, 8, 5)	0.31 [0.22, 0.40]	6.95 (<.0001)
rFFA2	24	40, –38, –22 (3, 7, 4)	0.28 [0.23, 0.34]	9.58 (<.0001)
Right OFA	24	29, –84, –23 (9, 7, 7)	0.17 [0.09, 0.26]	4.02 (.0004)
Right PHG	27	27, –55, –19 (3, 6, 4)	–0.23 [–0.26, –0.19]	–11.40 (<.0001)
lFFA1	27	–39, –59, –24 (4, 8, 5)	0.19 [0.12, 0.26]	5.00 (<.0001)
lFFA2	24	–40, –40, –24 (4, 7, 6)	0.18 [0.13, 0.23]	6.72 (<.0001)
Left OFA	23	–34, –80, –24 (9, 8, 5)	0.18 [0.08, 0.28]	3.55 (.0015)
Left PHG	27	–29, –53, –19 (4, 7, 4)	–0.27 [–0.31, –0.22]	–11.64 (<.0001)

**Table 3.** Correlations across Variables

	VET-LV	VET-NL	CFMT	Match-Bird	Match-Car/Plane
VET-LV	–				
VET-NL	.27	–			
CFMT	.11	.3	–		
Match-bird	.12	–.06	–.16	–	
Match-car/plane	–.35	<b>.55</b>	.1	.1	–
rFFA1	.04	.28	<b>–.46</b>	.17	.26
rFFA2	–.07	<b>.42</b>	–.1	.07	<b>.43</b>
lFFA1	<b>–.5</b>	.17	–.3	.14	<b>.58</b>
lFFA2	<b>–.68</b>	.05	–.18	–.11	<b>.53</b>
rOFA	–.25	.02	–.11	.03	.24
lOFA	–.3	–.08	–.31	.14	.1
rPHG	.05	.24	.21	.13	.16
lPHG	–.18	–.05	–.11	.21	.22

(Top) Zero-order correlations among behavioral variables: VET-LV, VET-NL, perceptual matching test with birds (Match-bird), average of perceptual matching test with cars and planes (Match-car/plane), and memory for faces (CFMT). (Bottom) Partial correlations between behavior and regional CT with participant age and global CT regressed out. (Note that regressing out age alone did not qualitatively change the results.) Significant correlations ( $p < .05$ ) are indicated in **bold**. We applied false discovery rate corrections (Benjamini & Hochberg, 1995) to each ROI for the three tests entered into multiple regression analyses—VET-LV, VET-NL, and CFMT (Table 5); the VET-NL correlation in rFFA2 failed to pass threshold.

ROI, thus leading to larger final ROIs relative to those initially defined. Functional voxels that were members of multiple initial ROIs were dropped from all final ROIs. This latter qualification avoided partial-volume effects with regard to functional region membership.

In addition to our functionally defined ROIs, we anatomically defined an additional four regions in the pre-central and frontal gyri to correspond to the regions where car expertise effects were reported in Gilaie-Dotan et al. (2012). We had no means to define this region functionally. The location and extent of these regions was fixed across all participants (see Table 4 legend).

To test whether CT varied as a function of ROI size and distance from the peak of face selectivity, we defined four additional clusters for bilateral FFA1 and FFA2 in each individual. First, we localized the peak face-selective voxel of each ROI based on the localizer scan. We computed mean CT from this peak voxel, in addition to the 4, 16, and 60 contiguous voxels around this peak, after the spread of face–object activation.

For all ROIs, we computed the partial correlation between the mean CT over all voxels with each VET factor, regressing out the other VET factor as well as global CT and age, because CT has been shown as highly sensitive to age (Shaw et al., 2008). Zero-order correlations and partial correlations for each ROI are presented in Table 3. All correlations between CT and behavioral performance were tested for bivariate outliers, which were denoted as points whose externally studentized residual was  $>3.5$  or  $<-3.5$ . Partial correlations are reported in Table 3.

To perform group-level statistical data analyses on CT maps, we used an advanced, HR, cortical matching approach (Frost & Goebel, 2012; Goebel, Hasson, Harel, Levy, & Malach, 2004; Goebel, Staedtler, Munk, & Muckli, 2002) to align brains using cortex curvature information (i.e., the gyral/sulcal folding patterns). Cortex-based alignment operates in several phases during which individual hemispheres are morphed into spheres providing a parameterizable surface suited for across-participant non-rigid alignment. Alignment proceeds iteratively following a coarse-to-fine matching strategy, moving from highly smoothed curvature maps to minimally smoothed maps (Frost & Goebel, 2012; Goebel, Esposito, & Formisano, 2006; Goebel et al., 2002, 2004).

Cortex-based alignment was used to compute average thickness maps across participants. Although CT measurements are performed in volume space in individual brains, they are performed in surface space for group analyses to benefit from cortical alignment.

During the segmentation procedure, all structural data sets were upsampled from the 1.0-mm isovoxel acquisition resolution to 0.5-mm isovoxel resolution using sinc interpolation. For whole-brain group analyses only, individual CT maps were smoothed by a factor of 2 times the size of the upsampled voxel, using 1-mm FWHM. These smoothed maps were subsequently used as input in a group correlation analysis.

We used a corrected two-tailed alpha of .05 for whole-brain analyses. These analyses seeking areas where CT correlated with VET-LV, VET-NL, and CFMT performance

failed to reveal significant clusters of activation. Whole-brain analyses are inherently less powerful than ROI analyses both because of correction for multiple comparisons and the greater variance expected when participants are compared in regions aligned according to gross anatomical rather than functional landmarks.

## CT

CT measurements in Brain Voyager QX are based on the Laplace method (Jones, Buchbinder, & Aharon, 2000). Three tissue classes are identified in the anatomical image based on a voxel's intensity value,  $i$ : cerebral spinal fluid ( $i < 75$ ), gray matter ( $75 \leq i \leq 125$ ), and white matter ( $i > 125$ ). For each gray matter voxel, a streamline is calculated—using a small step size of 0.1 and trilinear interpolation—by following a gradient in one direction and then the opposite direction to obtain a thickness measure for that gray matter voxel.

Measurement of CT of individual segmented cortical hemispheres is performed first in volume space but can be projected on the surface with the help of gradient maps. See Table 1 for descriptive statistics.

## RESULTS

### Relationship between Performance and CT

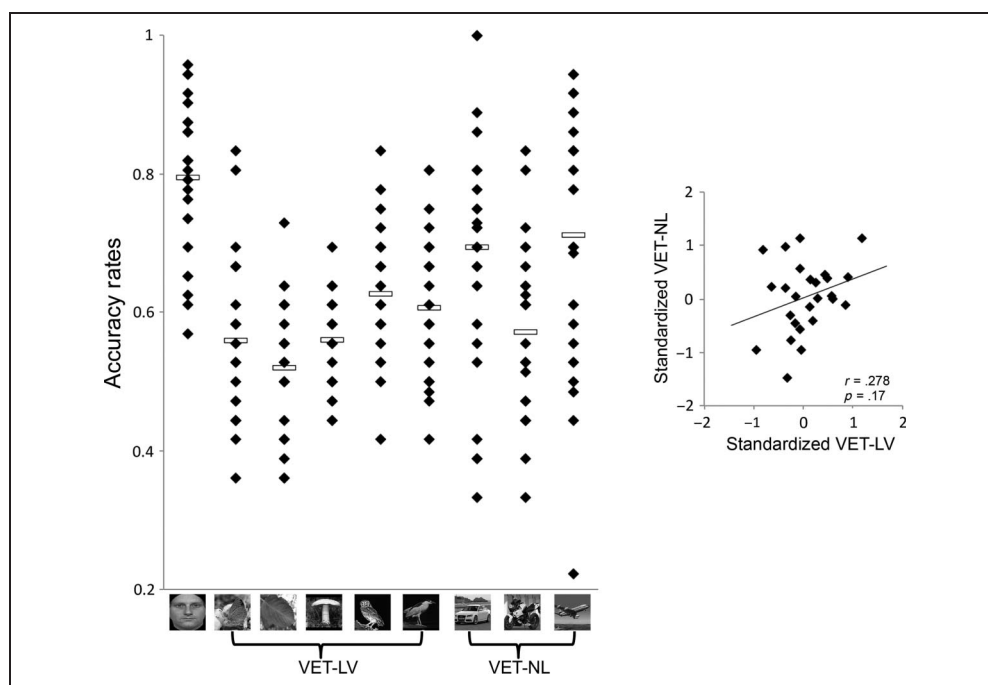
Just as living and nonliving performance scores were computed, so were living and nonliving SR scores. SR scores of experience for living and nonliving categories were significantly correlated ( $r = .48$ ), and the only significant correlation between SR and performance was that SR for nonliving objects negatively predicted VET-LV

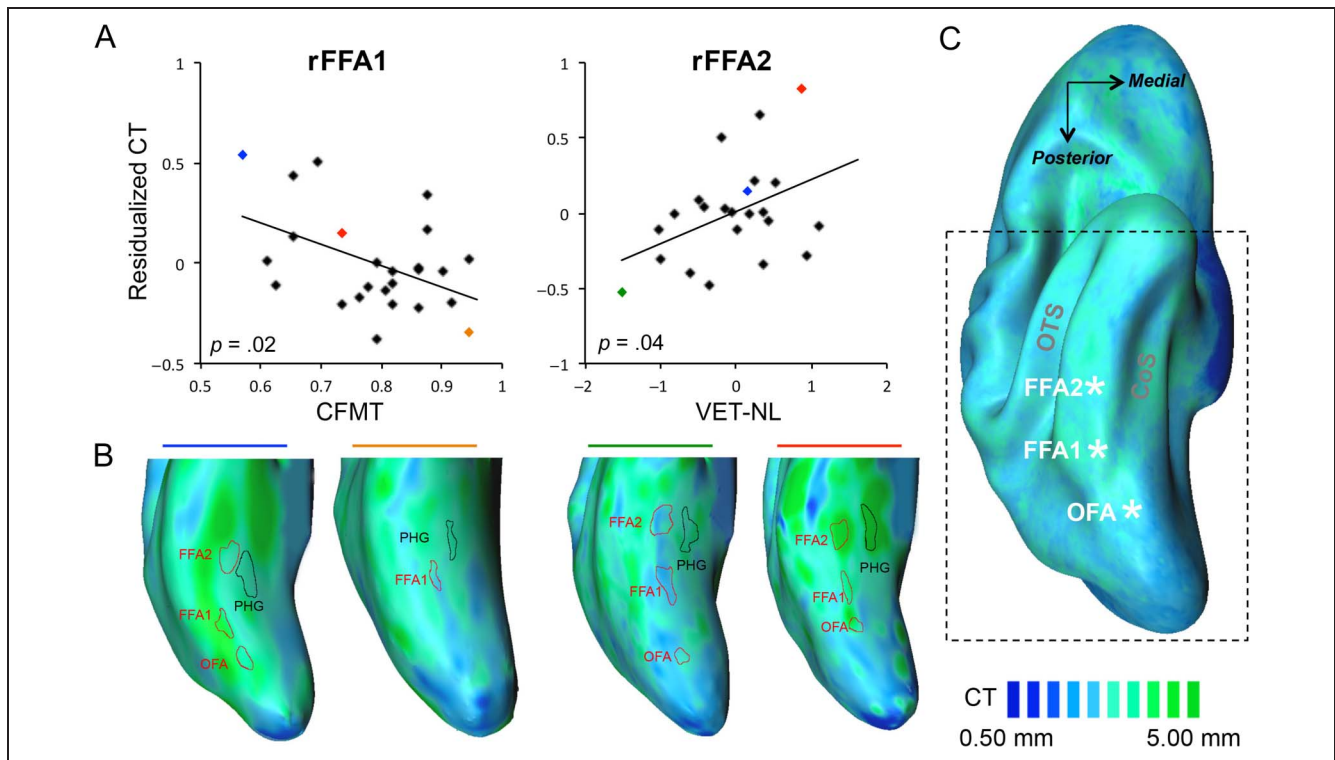
( $r = -.45$ ). These results are consistent with prior reports that SRs generally do a poor job predicting performance (McGugin, Gatenby, et al., 2012), probably because we have limited opportunity to compare our perceptual skills with those of others. In addition, SRs did not correlate significantly with CT in any ROI.

Table 3 provides correlations between our behavioral measures of performance with faces (CFMT) and living (VET-LV) and nonliving (VET-NL) object categories as well as the partial correlations that involve measures of CT in the various ROIs (we first regressed age and global CT out of the CT values within each ROI; see Figure 1 and Table 1 for CT averages and spreads). (Figure 1 shows the distribution of raw scores for CFMT and VET.) Performance with faces and nonface objects showed no significant correlation in this sample, although each measure was reliable (Cronbach's alpha: VET-LV = .89, VET-NL = .91) and showed considerable variability (Table 1).

Table 3 also presents the partial correlations between performance measures and CT across functional ROIs. The only significant effects were found in the FFAs (Figures 2 and 3). The only significant positive correlation for VET was in right FFA2 (rFFA2), where CT was related to VET-NL ( $r = .42$ ; Figure 2). To correspond to the VET scores, we grouped the matching performance for cars and planes ( $r = .57$ ), whereas birds was the only living category. Matching performance for cars/planes was correlated with VET-NL ( $r = .55$ ) and showed a similar positive correlation with CT in rFFA2 ( $r = .43$ ). Matching cars/planes produced the same positive correlation in the lFFAs, an effect that was not seen for VET scores (even when restricted to cars and planes, the correlations with the two lFFAs are both .24). We can only speculate that it is possible that the requirements of the matching

**Figure 1.** Dotplot depicting the behavioral performance in the CFMT (represented by the face stimulus) and the VET, grouped into VET-LV (butterflies, leaves, mushrooms, owls, and wading birds) and VET-NL (cars, motorcycles, and planes) categories. Each dot represents the accuracy of a given participant, and the horizontal bars represent the mean accuracy across participants for a given category. The scatterplot to the right shows the relationship between standardized measures of VET-LV and VET-NL.





**Figure 2.** (A) Scatterplots showing the significant partial correlations (regressing out participant age and global CT) between behavioral performance on faces (CFMT; left) and behavioral performance on nonliving object categories (VET-NL; right) with regional CT in rFFA1 and rFFA2, respectively. Colored points in the scatterplots correspond to the individual inflated hemispheres shown in B. (B) Four inflated right hemispheres, selected to demonstrate the most extreme (thickest or thinnest) FFA cortices as depicted by the scatterplots in A. Participant-specific maps of CT are overlaid on the corresponding inflated hemispheres, with functionally defined face-selective (FFA1/FFA2/OFA) and object-selective (PHG) ROIs outlined on top of the CT map. (C) Group-averaged CT map overlaid on the group-averaged inflated right hemisphere, with group-averaged coordinates for the center of rFFA1, rFFA2, and rOFA overlaid. Also labeled are the occipital temporal sulcus (OTS) and collateral sulcus (CoS). The dashed box represents the FOV for the hemispheres represented in B.

task tap better into left hemisphere representations, but this conjecture would have to be investigated.

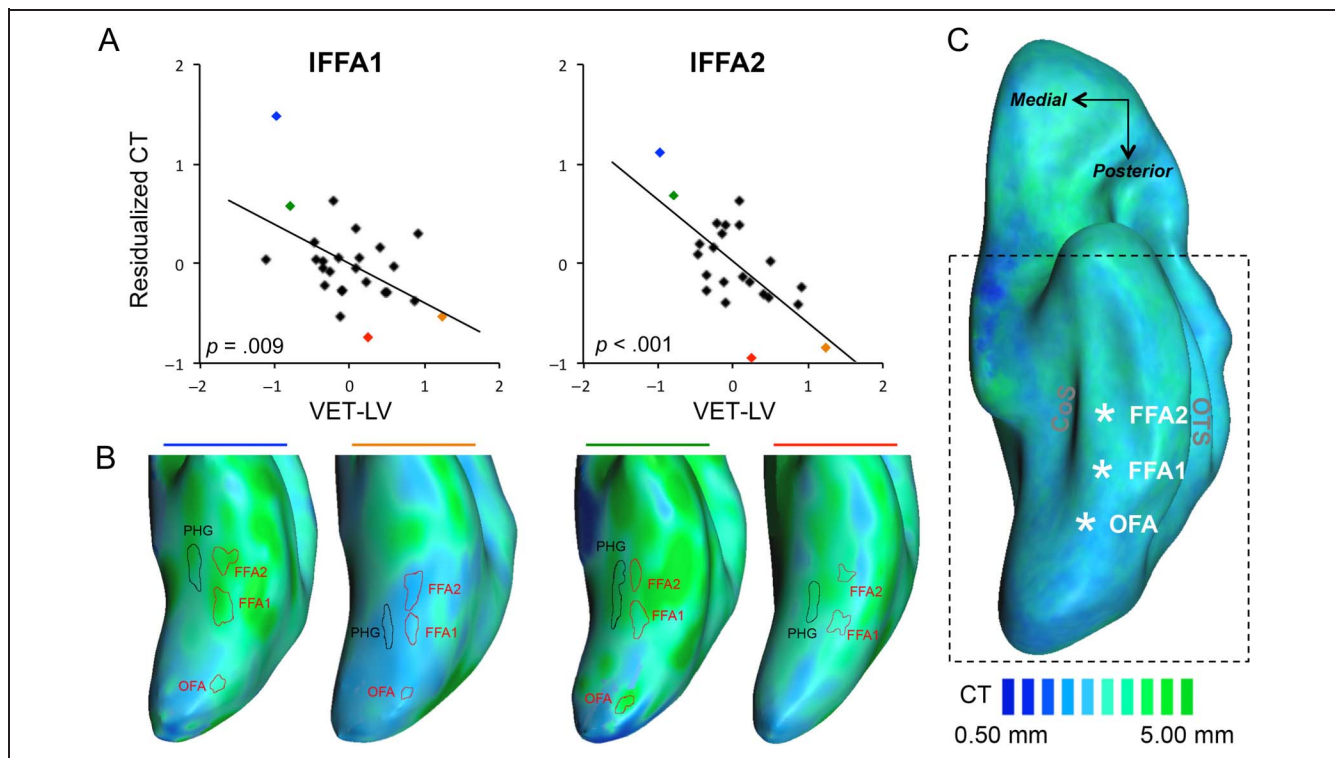
In contrast to these positive correlations for cars/planes, VET-LV showed significant negative correlations with CT in the two IFFA ROIs (Figure 3). Performance on the CFMT was negatively correlated with CT in rFFA1 (Figure 2). The matching task for birds did not correlate with CT in any area, although the only negative correlation was observed in the IFFA2 where the relationship with VET-LV was also most negative.

Interestingly, even when we restrict our analyses to consider thickness in the single maximally face-selective voxel, the pattern observed at the larger sized ROIs remains in IFFA1 ( $r_{\text{VET-LV}} = -.49$ ) and in rFFA2 ( $r_{\text{VET-LV}} = .39$ ). Other effects, however, were considerably reduced, including that of VET-LV in IFFA2 ( $r_{\text{VET-LV}} = -.30$ ) and of CFMT in rFFA1 ( $r_{\text{CFMT}} = -.22$ ). In addition to these ventral areas, we explicitly probed for frontal effects by defining four areas in the frontal and precentral gyri of all participants. These four ROIs were placed in regions showing CT effects of car expertise in prior work (Gilaie-Dotan et al., 2012). Only one region in the right superior frontal gyrus (rSFG) showed a positive correlation between behavioral performance (VET-LV) and regional CT ( $r = .41$ ; Table 4).

Finally, in contrast to our functionally and anatomically defined ROI results, whole-brain correlation analyses performed at the group level in average brain space did not reveal any significant effects between behavior and CT, even at a liberal threshold. Note that maps in Figures 2C and 3C depict average CT across all participants irrespective of behavior. Because of individual differences in CT, as well as error in cortical registration, these group maps do not reflect the full range of CT variability found in individual participants.

### Multiple Regressions on CT

Performance with faces and objects was not strongly related, and as such, it is possible that they account for different parts of the variance in CT. We conducted multiple regressions to assess how much variance in CT these variables could explain together in each ROI. All three predictors (CFMT, VET-LV, and VET-NL) were entered simultaneously in a multiple regression. The results for the four FFA ROIs are shown in Table 5, including the zero-order correlations (Table 3) for comparison with the partial correlations (note that they are not strictly



**Figure 3.** (A) Scatterplots showing the significant partial correlations (regressing out participant age and global CT) between behavioral performance on living object categories (VET-LV) with regional CT in IFFA1 (left) and IFFA2 (right). Colored points in the scatterplots correspond to colored bars above the individual inflated hemispheres represented in B. (B) Four inflated left hemispheres, selected to demonstrate the most extreme (thickest or thinnest) FFA cortices as depicted by the scatterplots in A. Participant-specific maps of CT are overlaid on the corresponding inflated hemispheres, with functionally defined face-selective (FFA1/FFA2/OFA) and object-selective (PHG) ROIs outlined on top of the CT map. (C) Group-averaged CT map overlaid on the group-averaged inflated left hemisphere, with group-averaged coordinates for the center of IFFA1, IFFA2, and IOFA overlaid. Also labeled are the occipital temporal sulcus (OTS) and collateral sulcus (CoS). The dashed box represents the FOV for the hemispheres represented in B.

**Table 4.** Partial Correlations across Behavioral Measures and Anatomically Defined Volumes Matched with Those in Which Car Expertise Effects on CT Were Observed in Gilaie-Dotan et al. (2012)

	VET-LV	VET-NL	CFMT
l aIFG	-.13	-.13	.24
r iPC	.19	.22	-.08
rMFG	-.07	-.22	-.11
rSFG	.41	.14	.28

Regions were defined based on the Montreal Neurological Institute coordinates (transformed to Talairach using MATLAB; The MathWorks, Natick, MA) and sizes reported in Gilaie-Dotan et al. (2012; Table 1) and were identical for all participants. Regions included left anterior inferior frontal gyrus (l aIFG): Tal ( $-42, 32, -1$ ), volume ( $174 \text{ mm}^3$ ); right inferior precentral sulcus (r iPC): Tal ( $47, 1, 2$ ), volume ( $520 \text{ mm}^3$ ); rSFG: Tal ( $12, 53, 10$ ), volume ( $47 \text{ mm}^3$ ); and right middle frontal gyrus (rMFG): Tal ( $29, 18, 37$ ), volume ( $27 \text{ mm}^3$ ). Behavioral measures included VET-LV, VET-NL, and faces (CFMT). We find a significant correlation between VET-LV and CT in the rSFG region, after we regress out the influence of global CT and age ( $r = .41, p = .028$ ).

speaking zero order because age and global CT were regressed out, but they do not take into account any of the other behavioral measures). Neither the full models nor the partial correlations were significant in the other non-FFA functionally defined ROIs.

These analyses allow us to ask how much unique variance is explained by each of the three measures. Although the simple correlations reveal that VET-NL was a significant predictor of CT only in rFFA2, when VET-LV and CFMT are partialled out, both the rFFA1 and IFFA1 also show the same positive correlation. This means that one or both of the other variables was suppressing this relation. We identified the suppressor by removing each variable in turn from the regressions. In the rFFA1, this suppressor variable was CFMT, and adding VET-LV had little influence on the VET-NL predictor. In the IFFA1, both of the other predictors were necessary for VET-NL to reach significance. In contrast, VET-LV remained a predictor in these multiple regressions, similar to when it was used as the sole behavioral predictor, in two areas: VET-LV accounted for unique variance (a negative correlation) in CT for both IFFA1 and IFFA2. Finally, there was unique variance in CT accounted for by the CFMT in both the rFFA1 and IFFA1.

**Table 5.** Multiple Regressions in the FFAs

<i>CT in rFFA1 (Age and Global CT Partialled Out)</i>						
<i>n = 25, RSq-adj = 33.2%</i>						
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>Partial r</i>	<i>Zero-order r</i>
Intercept	−0.0092	0.0402	−0.229	.821		
CFMT	−0.1540	0.0445	−3.46	.002	−.60*	−.46*
VET-LV	−0.0005	0.0741	−0.007	.995	.00	.04
VET-NL	0.1720	0.0655	2.63	.016	.50*	.28
<i>CT in rFFA2 (Age and Global CT Partialled Out)</i>						
<i>n = 26, RSq-adj = 13.7%</i>						
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>Partial r</i>	<i>Zero-order r</i>
Intercept	0.0033	0.0653	0.051	.960		
CFMT	−0.0846	0.0694	−1.22	.238	−.10	−.27
VET-LV	−0.0632	0.1261	−0.501	.622	−.07	−.11
VET-NL	0.2620	0.1058	2.48	.023	.42*	.49*
<i>CT in lFFA1 (Age and Global CT Partialled Out)</i>						
<i>n = 23, RSq-adj = 39.9%</i>						
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>Partial r</i>	<i>Zero-order r</i>
Intercept	1.1785	0.5287	2.23	.036		
CFMT	−1.4826	0.6598	−2.25	.035	−.43*	−.30
VET-LV	−0.4517	0.1261	−3.58	.002	−.61*	−.50*
VET-NL	0.2858	0.1119	2.56	.018	.48*	.17
<i>CT in lFFA2 (Age and Global CT Partialled Out)</i>						
<i>n = 23, RSq-adj = 42.5%</i>						
	<i>B</i>	<i>SE</i>	<i>t</i>	<i>p</i>	<i>Partial r</i>	<i>Zero-order r</i>
Intercept	0.4692	0.5963	0.787	.44		
CFMT	−0.5700	0.7413	−0.769	.45	−.17	−.18
VET-LV	−0.6236	0.1486	−4.2	.00	−.69*	−.68*
VET-NL	0.1380	0.1277	1.08	.29	.24	.05

### Relationship between Functional and Structural Effects of Expertise

The functional results for the present data set were presented in McGugin, Van Gulick, et al. (2014) and revealed a significant relationship between the BOLD response to cars relative to faces in both FFAs of both hemispheres, and when the BOLD response to birds was used as a baseline, there were significant effects of car expertise in rFFA2 and both lFFAs.

Our finding that behavior for different categories can be related to the CT in the same area in different ways

illustrates how difficult it would be to make predictions between such relative functional responses and CT measurements. The same ROI can yield many different responses for the same category depending on the task, whereas structural effects are stable and can reflect simultaneously the independent influence of many familiar categories.

Nonetheless, to test whether there was a link between the structural effects of CT and the functional BOLD-based effects of car expertise in McGugin, Van Gulick, et al. (2014), we correlated across participants the CT and the Michelson contrast ratios for cars (or faces) relative

**Table 6.** Correlations of Functional Selectivity for Faces with CT in Each ROI (at a Size of Four 3-mm<sup>3</sup> Voxels)

	<i>r</i> FFA1	<i>r</i> FFA2	<i>l</i> FFA1	<i>l</i> FFA2
fMRI C-B	.04	-.11	.13	-.06
fMRI F-B	-.33	-.12	-.15	.29

No correlations reached significance at  $p < .05$ .

to birds  $((\text{Car} - \text{Bird}) / (\text{Car} + \text{Bird}))$  and  $(\text{Face} - \text{Bird}) / (\text{Face} + \text{Bird}))$  in each FFA ROI (four standard resolution voxels). These functional responses were not significantly correlated with CT in any of the FFA ROIs (see Table 6). The largest effect size is observed in the relationship between CT and the face response in rFFA1 ( $r = -.33, p = .12$ ), which is in the same direction as the relation between CFMT and CT in this ROI. Future studies should consider functional responses to more object categories and the use of an unfamiliar object category as a baseline (so that effects can be investigated for each familiar category independently).

## DISCUSSION

We investigated how performance with objects and faces relates to CT in several individually defined functional ROIs. Our use of functionally defined ROIs afforded greater sensitivity over standard methods that are based on anatomical averaging. Gilaie-Dotan et al. (2012) also looked at individually defined FFAs and found no relation between CT and car expertise, although their sample was smaller (15 participants for rFFA). Several other differences could explain why we found effects and they did not; for example, we defined separate anterior and posterior FFAs and measured behavioral performance for more object categories. Our results suggest that, when the peaks of face selectivity are defined functionally, structural effects may be observed within very small regions centered on these peaks. We found a positive correlation between performance with nonliving objects and CT in FFA, whereas the relationship for faces and living objects with CT, when found, was negative. These CT results are generally consistent with past functional results in linking FFA specialization to nonface recognition, but the directions of the effects were unexpected. In addition, we found no evidence of a relation between BOLD responses to cars and faces (relative to birds) and CT in FFA ROIs, but future work should consider using a non-familiar category as baseline to look at the relation between each familiar category and CT measurements.

To our knowledge, this is the first study looking at CT separately in the anterior and posterior parts of human FFA (Weiner et al., 2010; Pinsk et al., 2009). We found that behavioral performance with faces has a greater contribution to CT in posterior parts of the FFA bilaterally. However, in none of the FFA ROIs did we find a relationship with face performance and not with object perfor-

mance. The current results present little evidence that any part of the FFA complex is selectively related to face but not object recognition.

Our results could be a function of the specific sample used in this study (male participants, selected on the basis of high or low SR of car expertise). In prior work, the relation between performance with faces and different object categories was found to be mediated by gender (McGugin, Richler, et al., 2012). In that work, women outperformed men on the VET-LV factor, whereas men performed better on the VET-NL factor (in this case, vehicles). When age and holistic processing of faces were partialled out, the unique variance explained by each VET factor was correlated with the CFMT, only for the gender-congruent category. Thus, it would be prudent not to generalize the present results to women: It is possible, albeit only a speculation, that the results in a sample of women might be a mirror image of those obtained here for men, with performance for living objects positively correlated with CT but performance for nonliving objects negatively correlated with CT. This may also be predicted on the basis of several studies reporting that women show an advantage on verbal tasks with living objects and men show an advantage for nonliving objects (Capitani, Laiacona, & Barbarotto, 1999; Laws, 1999; Laiacona, Barbarotto, & Capitani, 1998; McKenna & Parry, 1994).

Another consideration is that the functional definition of the FFA was based on a typical localizer that compared images of faces with images of manmade objects (tools, appliances, items of clothing, etc). Prior work has suggested that the location of the FFA is not impacted by the type of baseline (Berman et al., 2010), but we do not know of work that has compared localization based on a living-versus-nonliving comparison. We have no reason to believe that our results would vary if a different localizer was used, especially those effects that were essentially the same in a one-voxel ROI versus a 60-voxel ROI.

Our findings of a negative correlation between CT and face recognition converge with recent results showing that CT in the FFA was negatively correlated with learning performance on a face orientation judgment task (Bi et al., 2014). We found such a relationship in the rFFA1 (CT negatively correlated with face performance on the CFMT), whereas the previous work only found the effect in the lFFA (note that this learning study did not separate the two FFAs and used group-averaged ROI definitions). We also found that CT in both parts of the lFFA was negatively related to performance with living objects. Thus, our work considerably extends the Bi et al. finding to face recognition performance and suggests that such a negative correlation may not be specific to the lFFA or to performance with faces. It does not, however, provide insight into the biological mechanism that underlies this negative relationship. Negative correlations with performance have been attributed to synaptic pruning resulting

in the loss of nonpreferred cortical connections in favor of those that support frequently used skills (Gogtay et al., 2004; Sowell et al., 2004; Giedd et al., 1999). Another possible account is that the observed reduction in measured gray matter reflects an increase in myelination such that white matter growth encroaches upon what was classified as gray matter (Paus, 2005). This is consistent with recent results showing that fractional anisotropy of the white matter tracts from FFA to the anterior temporal lobe correlates with face recognition ability (Gomez et al., 2015). It is possible that, in our sample, those with thinner cortices also had larger white matter tracts connecting FFA to anterior areas.

By themselves, none of these accounts are sufficient to explain why the effect differs from the positive relationship obtained with nonliving objects. We obtained positive and negative relationships with performance in the same participants in the same areas, which may seem surprising, but the multiple regression analyses suggest that the different effects are independent. One possible explanation is that performance with these different categories reflects different ages of acquisition for experience individuating objects (arguably faces and perhaps also living objects, earlier than vehicles), with different mechanisms of plasticity operating at these different times. Face recognition could be learned early in life when pruning of large fiber tracts is taking place (Bourgeois, Jastreboff, & Rakic, 1989). In contrast, the recognition of vehicles could be learned much later in life and, as such, may show thickening of cortex as in learning of skills in adulthood (e.g., Mårtensson et al., 2012; Maguire, Woollett, & Spiers, 2006).

The relationships we show are not causal: Performance with a category would not cause CT, nor would CT cause performance, but rather, it is more plausible that experience with a category would cause both performance and CT. These are conjectures that should be explored in future research.

Critically, we find that nonface recognition can be predicted by CT in the FFA, an effect that cannot be accounted for by attention and providing further evidence that this region is important for nonface object processing. This should not be taken to suggest that other regions in the brain are not also involved in the ability to recognize objects and could also be shaped structurally by such experience. We found only limited replication of the prefrontal areas where CT correlated with car expertise in prior work, but unlike in FFA, we did not have individual functional ROIs to rely on. The effects of experience on brain structure may be variable and require methods that allow for spatial displacement of ROIs across individuals (see also Pinel et al., 2014). Finally, the structural effects of expertise have an interesting advantage over the more standard functional expertise effects: It could lead to a relatively faster accumulation of evidence across different laboratories, as a VET battery (free and available from authors) can be easily adminis-

tered to participants in the laboratory or online, before or after their participation in any study with a functional FFA localizer.

## Acknowledgments

We thank Benjamin J. Tamber-Rosenau and Jennifer J. Richler for assistance. This work was supported by the NSF (SBE-0542013), the Vanderbilt Vision Research Center (P30-EY008126), and the National Eye Institute (R01 EY013441-06A2).

Reprint requests should be sent to Rankin W. McGugin, Department of Psychology, Vanderbilt University, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37240-7817, or via e-mail: Rankin.McGugin@vanderbilt.edu.

## Note

1. The Grill-Spector et al. study used antique cars in the scanner when participants were modern car experts (see Bukach, Phillips, & Gauthier 2010), and the Brants et al. study did not provide behavioral evidence for the same qualitative changes in perception as the original study.

## REFERENCES

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.
- Berman, M. G., Park, J., Gonzalez, R., Polk, T. A., Gehrke, A., Knafila, S., et al. (2010). Evaluating functional localizers: The case of the FFA. *Neuroimage*, 50, 56–71.
- Bi, T., Chen, J., Zhou, T., He, Y., & Fang, F. (2014). Function and structure of human left fusiform cortex are closely associated with perceptual learning of faces. *Current Biology*, 24, 222–227.
- Bilalić, M., Langner, R., Ulrich, R., & Grodd, W. (2011). Many faces of expertise: Fusiform face area in chess experts and novices. *Journal of Neuroscience*, 31, 10206–10214.
- Bourgeois, J. P., Jastreboff, P. J., & Rakic, P. (1989). Synaptogenesis in visual cortex of normal and preterm monkeys: Evidence for intrinsic regulation of synaptic overproduction. *Proceedings of the National Academy of Sciences, U.S.A.*, 86, 4297–4301.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Brants, M., Wagemans, J., & Op de Beeck, H. (2011). Activation of fusiform face area by Greebles is related to face similarity but not expertise. *Journal of Cognitive Neuroscience*, 23, 3929–3958.
- Bukach, C. M., Phillips, S. W., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Performance and Psychophysics*, 72, 1865–1874.
- Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience*, 12, 679–690.
- Capitani, E., Laiacina, M., & Barbarotto, R. (1999). Gender affects word retrieval of certain categories in semantic fluency tasks. *Cortex*, 35, 273–278.
- Choi, Y., Shamosh, N., Cho, S., DeYoung, C., Lee, M., Lee, J., et al. (2008). Multiple bases of human intelligence revealed by cortical thickness and neural activation. *Journal of Neuroscience*, 28, 10323–10329.

- Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 94–107.
- Delon-Martin, C., Plailly, J., Fonlupt, P., Veyrac, A., & Royet, J.-P. (2013). Perfumers' expertise induces structural reorganization in olfactory brain regions. *Neuroimage*, 68, 55–62.
- Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, 44, 576–585.
- Foster, N. E. V., & Zatorre, R. J. (2010). Cortical structure predicts success in performing musical transformation judgments. *Neuroimage*, 53, 26–36.
- Frost, M. A., & Goebel, R. (2012). Measuring structural-functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage*, 59, 1369–1381.
- Garrido, L., Furl, N., Draganski, B., Weiskopf, N., Stevens, J., Tan, G. C. Y., et al. (2009). Voxel-based morphometry reveals reduced grey matter volume in the temporal cortex of developmental prosopagnosics. *Brain*, 132, 3443–3455.
- Gauthier, I., Curby, K. M., Skudlarski, P., & Epstein, R. A. (2005). Individual differences in FFA activity suggest independent processing at different spatial scales. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 222–234.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3, 191–197.
- Gauthier, I., & Tarr, M. J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, 37, 1673–1682.
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, 2, 861–863.
- Gilaie-Dotan, S., Harel, A., Bentin, S., Kanai, R., & Rees, G. (2012). Neuroanatomical correlates of visual car expertise. *Neuroimage*, 62, 147–153.
- Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with Brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human Brain Mapping*, 27, 392–401.
- Goebel, R., Hasson, U., Harel, M., Levy, I., & Malach, R. (2004). Statistical analyses across aligned cortical hemispheres reveal high-resolution population maps of human visual cortex. *NeuroImage Supplement*, 22, e975.
- Goebel, R., Staedtler, E., Munk, M. H. J., & Muckli, L. (2002). Cortex-based alignment using functional and structural constraints. *NeuroImage Supplement*, 16, 10533.
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences, U.S.A.*, 101, 8174–8179.
- Golestani, N., Paus, T., & Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron*, 35, 997–1010.
- Gomez, J., Pestilli, F., Witthoft, N., Golarai, G., Liberman, A., Poltoratski, S., et al. (2015). Functionally defined white matter reveals segregated pathways in human ventral temporal cortex associated with category-specific processing. *Neuron*, 85, 216–227.
- 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.
- Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences*, 12, 474–480.
- Harel, A., Gilaie-Dotan, S., Malach, R., & Bentin, S. (2010). Top-down engagement modulates the neural expressions of visual expertise. *Cerebral Cortex*, 20, 2304–2318.
- Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: When less is better than more. *Journal of Neuroscience*, 27, 13028–13032.
- Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: A two-site study. *Brain*, 129, 2562–2570.
- Jones, S. E., Buchbinder, B. R., & Aharon, I. (2000). Three-dimensional mapping of cortical thickness using Laplace's equation. *Human Brain Mapping*, 11, 12–32.
- Jung, R. E., Segall, J. M., Jeremy Bockholt, H., Flores, R. A., Smith, S. M., Chavez, R. S., et al. (2010). Neuroanatomy of creativity. *Human Brain Mapping*, 31, 398–409.
- Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, U.S.A.*, 107, 11163–11170.
- 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.
- Karama, S., Ad-Dab'bagh, Y., Haier, R. J., Deary, I. J., Lyttelton, O. C., Lepage, C., et al. (2009). Positive association between cognitive ability and cortical thickness in a representative US sample of healthy 6 to 18 year-olds. *Intelligence*, 37, 145–155.
- Laiacina, M., Barbarotto, R., & Capitani, E. (1998). Semantic category dissociations in naming: Is there a gender effect in Alzheimer's disease? *Neuropsychologia*, 36, 407–419.
- Laws, K. R. (1999). Gender affects naming latencies for living and nonliving things: Implications for familiarity. *Cortex*, 35, 729–733.
- Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus*, 16, 1091–1101.
- Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., et al. (2012). Growth of language-related brain areas after foreign language learning. *Neuroimage*, 63, 240–244.
- McGugin, R. W., Gatenby, C., Gore, J., & Gauthier, I. (2012). High-resolution imaging of expertise reveals reliable object selectivity in the FFA related to perceptual performance. *Proceedings of the National Academy of Sciences, U.S.A.*, 109, 17063–17068.
- McGugin, R. W., Newton, A. T., Gore, J., & Gauthier, I. (2014). Robust expertise effects in right FFA. *Journal of Neuropsychology*, 63, 135–144.
- McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Gauthier, I. (2012). The Vanderbilt Expertise Test reveals domain-general and domain-specific sex effects in object recognition. *Vision Research*, 69, 10–22.
- McGugin, R. W., Van Gulick, A. E., Tamber-Rosenau, B. J., Ross, D. A., & Gauthier, I. (2014). Expertise effects in face-selective areas are robust to clutter and diverted attention, but not to competition. *Cerebral Cortex*, 25, 2610–2622.
- McKenna, P., & Parry, R. (1994). Category specificity in the naming of natural and man-made objects: Normative data from adults and children. *Neuropsychological Rehabilitation*, 4, 255–281.
- Narr, K. L., Woods, R. P., Thompson, P. M., Szeszkó, P., Robinson, D., Dimitcheva, T., et al. (2007). Relationships between IQ

- and regional cortical gray matter thickness in healthy adults. *Cerebral Cortex*, 17, 2163–2171.
- Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences*, 9, 60–68.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging studies of attention: From modulation of sensory processing to top-down control. *Journal of Neuroscience*, 23, 3990–3998.
- Pinel, P., Lalanne, C., Bourgeron, T., Fauchereau, F., Poupon, C., Artiges, E., et al. (2014). Genetic and environmental influences on the visual word form and fusiform face areas. *Cerebral Cortex*, 25, 2478–2493.
- Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., et al. (2009). Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *Journal of Neurophysiology*, 101, 2581–2600.
- Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage*, 19, 473–481.
- Rossion, B., Kung, C.-C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, 101, 14521–14526.
- Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., et al. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nature Neuroscience*, 8, 1241–1247.
- Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience*, 14, 28–30.
- Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., et al. (2006). Intellectual ability and cortical development in children and adolescents. *Nature*, 440, 676–679.
- Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., et al. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *Journal of Neuroscience*, 28, 3586–3594.
- Shomstein, S., & Yantis, S. (2006). Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *Journal of Neuroscience*, 26, 435–439.
- Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience*, 24, 8223–8231.
- Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, 3, 764–769.
- Weiner, K. S., Sayres, R., Vinberg, J., & Grill-Spector, K. (2010). fMRI-adaptation and category selectivity in human ventral temporal cortex: Regional differences across time scales. *Journal of Neurophysiology*, 103, 3349–3365.
- Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., et al. (2014). The mid-fusiform sulcus: A landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *Neuroimage*, 84, 453–465.
- Wonderlick, J. S., Ziegler, D. A., Hosseini-Varnamkhasti, P., Locascio, J. J., Bakkour, A., van der Kouwe, A., et al. (2009). Reliability of MRI-derived cortical and subcortical morphometric measures: Effects of pulse sequence, voxel geometry, and parallel imaging. *Neuroimage*, 44, 1324–1333.
- Wong, P. C. M., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadegh, A., Parrish, T. B., et al. (2008). Volume of left Heschl's gyrus and linguistic pitch learning. *Cerebral Cortex*, 18, 828–836.
- Xu, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex*, 15, 1234–1242.