Thickness of Deep Layers in the Fusiform Face Area Predicts Face Recognition

Rankin W. McGugin1, Allen T. Newton2,3,4, Benjamin Tamber-Rosenau5, Andrew Tomarken1, and Isabel Gauthier1

Abstract

People with superior face recognition have relatively thin cortex in face-selective brain areas, whereas those with superior vehicle recognition have relatively thick cortex in the same areas. We suggest that these opposite correlations reflect distinct mechanisms influencing cortical thickness (CT) as abilities are acquired at different points in development. We explore a new prediction regarding the specificity of these effects through the depth of the cortex: that face recognition selectively and negatively correlates with thickness of the deepest laminar subdivision in face-selective areas. With ultrahigh resolution MRI at 7T, we estimated the thickness of three laminar subdivisions, which we term “MR layers,” in the right fusiform face area (FFA) in 14 adult male humans. Face recognition was negatively associated with the thickness of deep MR layers, whereas vehicle recognition was positively related to the thickness of all layers. Regression model comparisons provided overwhelming support for a model specifying that the magnitude of the association between face recognition and CT differs across MR layers (deep vs. superficial/middle) whereas the magnitude of the association between vehicle recognition and CT is invariant across layers. The total CT of right FFA accounted for 69% of the variance in face recognition, and thickness of the deep layer alone accounted for 84% of this variance. Our findings demonstrate the functional validity of MR laminar estimates in FFA. Studying the structural basis of individual differences for multiple abilities in the same cortical area can reveal effects of distinct mechanisms that are not apparent when studying average variation or development.

INTRODUCTION

Brain volumetrics are behaviorally relevant, task-independent biomarkers that can be studied in terms of their relation to visual abilities. Those abilities depend on both genetic and environmental factors, as well as on the developmental state of the brain. Importantly, many structural biomarkers can be measured independent of behavior and are therefore more likely to reflect stable variability across individuals rather than performance during a specific task. The relation between brain structure and behavioral abilities can provide a useful window into the underlying forces that culminate into individual differences.

The current work builds on our previous finding that differences in total cortical thickness (CT) of the same brain area in the same people predict different abilities in opposite directions: Vehicle recognition (VR) was associated with a relatively thicker fusiform face area (FFA), whereas face recognition (FR) was associated with a relatively thinner FFA (McGugin, Van Gulick, & Gauthier, 2016). The current study extends that previous work, using ultrahigh resolution (UHR) imaging to probe laminar thickness differences underlying the differences in total CT, and their relationship to behavioral abilities.

When we reported the surprising pattern of FFA CT variations across individuals, we suggested that those opposing results (thinner CT associated with better FR, thicker CT associated with better VR, all within the same region of the brain) could be explained by differences in the developmental state (plasticity) of the brain at the time these visual skills were learned (McGugin et al., 2016). This was predicted because individuation for exemplars in the categories used, faces and vehicles, were likely learned at different times in development. It seemed plausible that learning occurring at different times in development would have different effects on underlying brain structure, with some of these effects being selective to specific cortical layers. In other words, our predictions are based on intuitions relevant to two different fields: the development of behavioral recognition abilities across domains and the development of the associated cortical underpinnings. These are two areas of study that are difficult to study, but we based our conjectures on the following findings.

First, we predicted that observed differences in brain correlates of FR and VR in adults may reflect differential ages of acquisition for these abilities (McGugin et al., 2016). There is growing evidence that faces dominate the visual diet of infants. Three-month-olds see faces for 25% of their waking time, an amount reduced to 10% by 12 months (Sugden & Moulson, 2019; Fausey, Jayaraman, & Smith,
2016), and early face exposure has been found to be critical to the development of face selectivity (Arcaro, Schade, Vincent, Ponce, & Livingstone, 2017). The age at which infants learn to individuate objects such as vehicles is less clear (Clerkin, Hart, Rehg, Yu, & Smith, 2017). Although face selectivity continues to develop during adolescence (Gomez et al., 2017; Scherf, Thomas, Doyle, & Behrmann, 2014) and FR continues to improve in adults (Germine, Duchaine, & Nakayama, 2011), it is reasonable to assume that FR also starts earlier than VR.

Second, we conjectured that if the earlier acquisition of FR (relative to VR) was responsible for a pattern where thinner cortex related to better performance, then the underlying structural biomarkers for the thinner cortex may be related to the early developmental pruning of connections critical to the acquisition of FR. The amygdala is critical to the development of FR (Schultz, 2005) and is connected to the occipital lobe via the inferior longitudinal fasciculus. Research on the connectivity between the inferior temporal (IT) areas TE and TEO (the monkey analog of the FFA) and the limbic system has revealed projections in the infant monkey that are pruned in the adult (Webster, Ungerleider, & Bachevalier, 1991). Such connections from IT to limbic areas are labeled predominantly in infragranual (deep) layers of IT (Stefanacci & Amaral, 2000). Accordingly, if this variability in the pruning of the inferior longitudinal fasciculus connections from IT to the amygdala is relevant to FR in humans, it should be most reflected in the thickness of deep cortical layers. In contrast, we expect the positive correlation between VR and variability in FFA’s CT to reflect mechanisms that support learning-dependent structural plasticity in adults (Wenger et al., 2012). These mechanisms are not well understood but are proposed to include processes such as gliogenesis and angiogenesis that accompany synaptogenesis (Zatorre, Fields, & Johansen-Berg, 2012). We have no reason to expect these mechanisms to be confined to deep layers (but acknowledge no evidence against laminar specificity).

It is important to note that multiple mechanisms may underlie any observed differences in structure that are associated with development. For example, variation in the pruning of connections could occur as described in the predictions that motivated the current work, affecting the thickness of different layers. However, the predicted pattern of laminar-specific correlations could be consistent with pruning but could also reflect other mechanisms. In particular, myelination increases during development and the effects of pruning and myelination are difficult to distinguish, as myelination could shift the apparent gray matter (GM)—white matter (WM) boundary and result in thinner cortical measurements, an effect that would be attributed to deep cortical layers (Sowell et al., 2004). It is fortuitous that both of these plausible mechanisms predict that the earlier development of FR would result in variance across individuals for FR to be specifically related to the thickness of the deep layers in our measurements.

In summary, here we test whether changes in CT that are associated with behavior (McGugin et al., 2016) are predominantly driven by changes in deep layers of the cortex, suggestive of developmental underpinnings. We additionally demonstrate that UHR MRI can be used to image laminar thickness within the cortex, providing a potentially useful biomarker for probing the structure–function relationship of the brain in vivo. In testing our predictions, we contrasted two models (Figure 1). The first represents a null hypothesis with regard to laminar
specificity: It simply assumes that the effects observed in overall CT in McGugin et al. (2016) will replicate here but will be the same across all layers (Model 1). In contrast, Model 2 instantiates our predictions based on differential early development and posits a selective linkage between the deep MR layer and variability in FR.

METHODS

Participants

Fourteen right-handed men aged 19–29 years (mean = 22.46 years, SD = 3.60 years) were recruited from the Vanderbilt University community based on previous participation. Only men were recruited to limit variability in brain anatomy. Our sample size was limited by the short period of time between when we developed scan protocols with sufficient resolution and before a major scanner upgrade. We recruited individuals with experience in the scanner and who, as a group, offered a wide range of performance with faces. One participant was excluded because of excessive head motion. Written informed consent was obtained from all participants in accordance with guidelines of the Vanderbilt University institutional review board and Vanderbilt University Medical Center. All participants reported normal or corrected-to-normal visual acuity and received monetary compensation for participation.

Behavioral Testing

Participants performed a battery of matching and recognition memory tests outside the scanner. First, they completed the extended Cambridge Face Memory Test (Russell, Duchaine, & Nakayama, 2009) in which they studied several views of unfamiliar male faces and were tested in forced-choice trials of three faces on which they selected the face that matched a studied face. Second, participants completed various sections of the Vanderbilt Expertise Test (VET; McGugin, Richler, Herzmann, Speegle, & Gauthier, 2012). This included two face tests (VET-Males, VET-Females), two vehicle tests (VET-Car, VET-Plane), and two animal tests (VET-Bird, VET-Butterfly). In each case, participants studied six identities, models, or species for as long as needed, followed by three-alternative forced-choice test trials in which they indicated which item was studied. Third, participants completed the Vanderbilt Face Matching Test (Sunday, Lee, & Gauthier, 2018). On each trial, participants studied two faces for 4000 msec and then were shown a three-alternative forced choice and indicated which face was a new image of a face studied on that trial. Matching faces were different images from the original presentation. Fourth, participants completed sequential matching tests using cars, planes, birds, and butterflies (McGugin, Gatenby, Gore, & Gauthier, 2012). On each trial, an image appeared for 1000 msec, followed by a 500-msec mask, and then a second image. Participants judged whether the two images showed cars/planes of the same make and model regardless of year or birds/butterflies of the same species. See Table 1 for mean test performance, reliability of each test, and consistency across tests for each category. Aggregate face, vehicle, and animal indices were calculated based in each case on standardized performance for the four relevant tests.

MRI Acquisition

All participants were imaged using a Philips Achieva whole-body 7T MRI scanner with a quadrature, head-only transmit coil and a 32-channel receive coil array. Imaging proceeded in three stages: whole-brain structural imaging, functional localization, and UHR susceptibility-weighted imaging (SWI).

Whole-brain Structural Imaging Acquisition

Whole-brain structural imaging consisted of a sagittal whole-brain 3D T1 weighted MPRAGE image gathered at 1-mm isotropic resolution and resampled into axial and coronal orientations for accurate localization of SWI acquisitions. This was used to identify anatomic

Table 1. Behavioral Results across 12 Independent Tests

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD)</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Faces</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFMT+</td>
<td>0.65 (0.13)</td>
<td>0.88</td>
</tr>
<tr>
<td>VET-Male (Acc)</td>
<td>0.83 (0.17)</td>
<td>0.78</td>
</tr>
<tr>
<td>VET-Female (Acc)</td>
<td>0.83 (0.12)</td>
<td>0.86</td>
</tr>
<tr>
<td>VFMT (Acc)</td>
<td>0.6 (0.07)</td>
<td>0.86</td>
</tr>
<tr>
<td><strong>Vehicles</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VET-Car (Acc)</td>
<td>0.73 (0.18)</td>
<td>0.93</td>
</tr>
<tr>
<td>VET-Plane (Acc)</td>
<td>0.76 (0.12)</td>
<td>0.93</td>
</tr>
<tr>
<td>Matching-Car (d’)</td>
<td>1.79 (1.00)</td>
<td>0.75</td>
</tr>
<tr>
<td>Matching-Plane (d’)</td>
<td>1.57 (0.77)</td>
<td>0.86</td>
</tr>
<tr>
<td><strong>Animals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VET-Bird (Acc)</td>
<td>0.68 (0.17)</td>
<td>0.80</td>
</tr>
<tr>
<td>VET-Butterfly (Acc)</td>
<td>0.56 (0.17)</td>
<td>0.83</td>
</tr>
<tr>
<td>Matching-Bird (d’)</td>
<td>1.46 (0.39)</td>
<td>0.92</td>
</tr>
<tr>
<td>Matching-Butterfly (d’)</td>
<td>1.65 (0.52)</td>
<td>0.82</td>
</tr>
</tbody>
</table>

For one participant, performance on VET-Male was near chance and an outlier relative to other participants and his own other scores, so this data point was dropped. The ICC4 (with 95% HPD interval) was .89 (.77, .96) for the aggregate of face tests, .83 (.66, .95) for the aggregate vehicle tests, and .85 (.70, .96) for the aggregate animal tests. CFMT+ = extended Cambridge Face Memory Test; VET = Vanderbilt Expertise Test; VFMT = Vanderbilt Face Matching Test.
landmarks important to the planning of the subsequent UHR SWI acquisitions.

fMRI Acquisition, Stimuli, Design, and Analyses

Functional localization of the FFA was accomplished via isotropic whole-brain fMRI data. Images were acquired with minimal distortion using a multishot 3D gradient-echo PRESTO EPI pulse sequence (Versluis et al., 2010): repetition time/echo time = 21.95/28.13 msec, flip = 12°, field of view = 211.4 mm, number of slices = 30, voxel size = 2.5 mm isotropic, EPI factor = 11, water/fat shift = 6.705 pix, number of dynamics = 160, time/dynamic = 2.02 sec.

We presented all stimuli with a Macintosh computer (Apple, Inc.) running MATLAB 2014a (The MathWorks) using Psychophysics Toolbox (Brainard, 1997). Stimuli were displayed on a rear projection screen using an Avotec projector. Immediately following the HR structural scan, each participant completed one to two runs (depending on time) of a functional localizer scan (160-dynamics/run). We used grayscale images (36 faces, 36 objects) in a 1-back detection task across 20 alternating blocks of face and object images. Each image was presented for 1 sec, with 16 images per category block. The HR T1-weighted structural scans were transformed into the standard sagittal view with 1-mm isometric voxels to facilitate coregistration of functional and anatomical data sets. Functional data were analyzed using Brain Voyager (www.brainvoyager.com) and in-house MATLAB scripts. Preprocessing in Brain Voyager included registration to the structural scan, slice scan time correction (cubic spline), 3D motion correction (sinc interpolation), and temporal filtering (high-pass criterion of two cycles per run) with linear trend removal. No spatial smoothing was applied. ROIs were defined using the Face > Object contrast from the face localizer scan. We localized an ROI that responded more to faces than objects in the right middle fusiform gyrus (FFA2) of all participants.

Right FFA2s were localized on the 3 mm³ (interpolated) statistical maps from the localizer scans as the peak face-selective voxel, then grown to incorporate contiguous 3-mm³ voxels passing statistical threshold within a four-voxel or 108 mm³ fixed volume around the peak of face selectivity. Solely for the purpose of comparing the location of our rFFA2s with those reported in the literature and demonstrating homology of the rFFA2 across participant using slice-selective Cartesian gradient-echo acquisitions. Both real and imaginary images were recorded. Measurement of individual differences in CT presents additional challenges over characterizing average laminar profiles because of the various sources of measurement error that can confound interindividual variability. Thus, care was taken to plan image acquisitions such that the frequency encoding direction was perpendicular to the ventral surface of the temporal lobe underneath the right FFA to minimize differences in partial volume effects across layers from participant to participant. Before analysis, UHR data were reconstructed at 0.1875 mm in-plane resolution.

SWI Image Acquisition, Processing, and Averaging

A minimum of six UHR T²*-weighted images were acquired as separate and independent acquisitions in each participant using slice-selective Cartesian gradient-echo acquisitions. Both real and imaginary images were recorded. Measurement of individual differences in CT presents additional challenges over characterizing average laminar profiles because of the various sources of measurement error that can confound interindividual variability. Thus, care was taken to plan image acquisitions such that the frequency encoding direction was perpendicular to the ventral surface of the temporal lobe underneath the right FFA to minimize differences in partial volume effects across layers from participant to participant. Before analysis, UHR data were reconstructed at 0.1875 mm in-plane resolution.

SWI imaging parameters were as follows: field of view = 240 × 180.194 × 21.9 mm, yielding voxel dimensions of 0.194 × 0.194 × 1.00 mm, 20 slices, 0.1 mm gap, “shortest” (878.8 ± 8.29 msec) repetition time, “shortest” (27.5 ± 0.31 msec) echo time, 27.26 pix water/fat shift, 55° flip angle, flow compensation, 9 min 11 ± 5.2 sec total duration. Images were acquired with a 1-D phase navigator to correct for phase errors arising from respiration during acquisition, as has been previously shown to be effective in very high resolution SWI imaging (Versluis et al., 2010).

Real and imaginary images were used to calculate magnitude and phase images, which were then processed to create susceptibility weighted scans (Haacke, Xu, Cheng, & Reichenbach, 2004). Within each participant, each of the six resulting SWI scans were coregistered using SPM12 (www.fil.ion.ucl.ac.uk/spm/) and averaged in MATLAB. For those participants who had more than six T²*-weighted scans acquired, the optimal six were identified for analysis by calculating the normalized mutual information (NMI; Studholme, Hill, & Hawkes, 1999; Viola & Wells, 1997; Wells, Viola, Atsumi, Nakajima, & Kikinis, 1996) between all pairs of acquired and coregistered SWI scans and ranking each according to the sum of its pairwise NMI values. The six scans with the highest summed pairwise NMI were kept for further analysis. This ensured that the six scans most similar to each other and likely containing the least substantial run-dependent artifacts or misalignments were selected for the purpose of averaging in later stages of the analysis.

Image Registration

Standard resolution functional localizer scans were registered to high-resolution (1 mm isotropic) structural scans in BrainVoyager (Figure 2A). Because we sought to avoid unnecessary interpolations of the UHR images, all intensity analyses were performed directly on the UHR images. We registered the high-resolution structural image and the functionally defined rFFA2 to the UHR slices. Specifically, high-resolution structural images and rFFA2 masks were converted from BrainVoyager’s proprietary image formats to NIfTI-1 images using the NeuroElf
toolbox for MATLAB. Rigid body transformation matrices of the high-resolution structural images to the UHR images were calculated via SPM’s SPM_coreg() coregistration function, and these matrices were used to reslice the rFFA2 masks into the UHR space. The 108 mm$^3$ rFFA2 masks traversed four to seven UHR slices, depending on the participant (mean = 5.57, SD = 0.85). For each UHR slice, 16-bit TIFF-format two-dimensional slice images were saved using MATLAB.

**Image Segmentation and Trace Identification**

Using Adobe Photoshop, the GM-WM (deep) and GM-cerebrospinal fluid (CSF; superficial) cortical boundaries were manually drawn by two image analysts trained to recognize differences in brain structures but blind to the experimental goals (Figure 2B). The interrater reliability of total rFFA2 thickness was very high (intraclass correlation [ICC] = .98). In addition to segmenting cortex from other tissue, blood vessels and image artifacts were manually marked as well. The resulting segmented data images were masked with the ROI images to create a series of ROI-specific cortical segmentation images. These segmented ROI images were next used to trace within-slice line segments through the thickness of cortex, along which we sampled the cortical signal intensity. Before tracing, each boundary was smoothed to avoid single-pixel concavities or convexities. Specifically, for each voxel that was part of the boundary, the three contiguous nearest neighbor boundary voxels in either direction (for a total of seven) were identified and a line segment fit to those voxels using linear regression. The set of such line segments was aggregated as the boundary was traversed and skeletonized (to removed thickenings) and cleaned of spurs using the MATLAB function bwmorph().

Because of the convoluted shape of the cortical sheath, points on the deep and superficial boundaries of cortex do not have a one-to-one correspondence. Using custom MATLAB code, we selected lines originating in each superficial boundary voxel and terminating in the nearest deep boundary voxel within that ROI in that slice (Figure 2C). We conducted the same procedure in the reverse direction as well, originating in each deep boundary voxel and terminating in each superficial boundary voxel. The resulting traces were subject to five constraints: First, traces could not intersect cortical boundaries (aside from at trace origination or termination) to ensure that all samples were restricted to cortical tissue even in areas of high cortical convolution. Second, no trace could cross a previously identified trace originating from the same boundary to avoid overrepresenting certain pieces of the cortical mantle. Third, the angle between the trace and boundary from which it originated could not be more than 20° off perpendicular to ensure that traces were approximately perpendicular to the cortical boundary. Fourth, any trace that was 3 SDs or more longer or shorter than the mean of all traces in a given slice was removed to control for outlier paths. Fifth, no trace could cross an area marked for exclusion, such as a vein, because MRI intensity values are not valid indicators of GM signal in veins. For any trace that violated one of these constraints, a new trace...
was drawn from the originating voxel to the next closest voxel in the other boundary.

Intensity Sampling and MR Layer Identification
First, average CT was defined in each ROI in each slice per participant as the average distance between the GM-CSF border and GM-WM border. Next, three cortical depths were identified by the points at which GM intensity trends changed (Figure 2D). Specifically, to estimate these depths, we fit a fourth-order polynomial to the set of intensity samples from each ROI in each slice. GM intensity was sampled at 500 equally spaced points along the length of each trace. This ensured that voxel intensities were sampled even when a trace only crossed a corner of a voxel. The sampled intensities were linearly interpolated to 1/100th mm intervals. The resulting intensities were thresholded since pixels above or below a certain intensity are assumed to reflect partial-volume or whole-pixel capture of non-GM. Specifically, any intensity values 3 SDs or more from the mean (over the entire set of trace intensities) were not included. Subsequently, intensity data were de-trended.

In all slices, inflection points were identified as the depths at which the second derivative of the polynomial was equal to zero. Inflection points outside the range of depths that fell within cortex were discarded, and the remaining points were taken as estimates of the boundaries between cortical layers. Thus, we operationalized three MR cortical layers as the distances between (1) GM-CSF border and the first inflection point (superficial cortical depth), (2) first and second inflection points (middle cortical depth), and (3) second inflection point and the GM-WM border (deep cortical depth). We hypothesize that these MR layers are related to, although they are not a direct measure of traditional cortical layers defined based on histological properties (Lifshits et al., 2018).

In 2 of 69 slices analyzed, the variation of intensity through the cortex was insufficient for accurately fitting a polynomial to the underlying data. These slices were omitted from further analyses.

Transcortical Microvasculature
Transcortical microvasculature was segmented and mapped using a contour-based local minimization routine. Having manually identified the deep and superficial boundaries of the functionally defined rFFA2, intermediary contours parallel to the cortical surface were drawn throughout the CT. Transcortical vessels were segmented by mapping local minima along each contour, and the vascular area was quantified as the percentage of voxels within the rFFA2 labeled as belonging to transcortical vessels. The measure for each participant was an average from at least four and up to seven slices, with a weighted average ICC of .80.

RESULTS
Reliability of CT Measurements
Slices
We estimated the reliability of cortical measurements across slices (after centering them on the peak of face selectivity) using the correlation of thickness estimates for each slice relative to all the other slices in each participant. For total regional CT, which was based on manual tracings of high-contrast boundaries, correlations were higher than .8 for all slices. As a result, total CT measurements were averaged over all available slices. Interrater reliability of total CT was calculated using ICCs for distance from GM-WM boundary to GM-CSF boundary, averaged across all slices for the two analysts.

Thickness measurements of the middle MR layer were based on relatively lower contrast boundaries identified using an automated algorithm (see below). Correlations reached .5 for anterior slices, but slices that were posterior relative to the peak of face selectivity showed little to no evidence of reliability. Inspecting individual images, this proved to be because FFA2 ROIs were located closer to the posterior slices of the UHR stack of 20 images for most participants. Our measurements were made after coregistration of six UHR scans to increase contrast; thus, the final slices of the stack were more likely to show coregistration artifacts due to motion or respiration. We therefore used only the two to four reliable anterior slices for each participant and dropped the two to three most posterior and unreliable slices, for all measurements involving an internal laminar boundary (thickness estimates for deep, superficial and middle MR layers).

Average Thickness Measures
The WM-GM boundary and the GM-CSF boundary were traced manually by two image analysts. Across participants, the mean total CT of rFFA2 (distance from the WM-GM boundary to the GM-CSF boundary) was consistent with previous work. Between the two analysts blind to behavioral measurements, interrater reliability of total rFFA2 thickness was very high: ICC = .98. MR-defined laminar subdivisions are defined by two internal boundaries identified using an automated procedure (see Methods). The representation of each MR laminar subdivision proportional to total CT was also consistent with previous work (Figure 2D). Across participants, the mean total CT of rFFA2 (distance from the WM-GM boundary to the GM-CSF boundary) was in the range reported in extrastriate cortex (Fischl & Dale, 2000): average rFFA2 GM thickness = 3.37 mm, with a range between 2.50 and 4.67 mm. Furthermore, the representation of each laminar subdivision proportional to total CT was consistent with MR microscopy in area V4 in nonhuman primates (Chen, Wang, Gore, & Roe, 2012): superficial MR layer (1.07 mm, range = 0.74–1.61 mm, 31%), middle MR layer (1.04 mm, range = 0.64–1.49 mm, 51%), deep MR layer (1.26 mm,
range = 0.56–2.72 mm, 38%). As an index of consistency across participants, ICCs for the three laminar subdivisions were: superficial (.77), middle (.73), and deep (.80).

**Association between Behavioral Measures and Total rFFA2 CT**

Each behavioral measure was computed as the average of the standardized values of four tests. ICCs were greater than .84 in all cases. We replicated previous results showing that total FFA CT is negatively correlated with FR ($r = -0.83, p < .001$) and positively correlated with the recognition of vehicles ($r = 0.70, p < .001$; Table 2), with the two correlations significantly different ($t = 8.86, p < .001$). Face and vehicle performance were not significantly related ($r = -0.33, p = .27$) and partial correlations with rFFA2 controlling for performance on the other categories were comparable in value and statistical significance to the zero-order $r$s (Table 2; Figure 3A). We replicated these correlational results using the univariate observed–imposed bootstrap (e.g., Beasley et al., 2007). Multiple regression indicated that both face and vehicle scores independently predicted rFFA2 thickness (both

**Table 2. Pearson’s Correlations between Behavioral Performance and CT and Laminar Depth Measurements**

<table>
<thead>
<tr>
<th></th>
<th>Total Thickness</th>
<th>Superficial Cortical Subdivision</th>
<th>Middle Cortical Subdivision</th>
<th>Deep Cortical Subdivision</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Zero-order Correlations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faces</td>
<td>$-0.83$</td>
<td>$-0.38$</td>
<td>$-0.28$</td>
<td>$-0.76$</td>
</tr>
<tr>
<td>Vehicles</td>
<td>$0.70$</td>
<td>$0.44$</td>
<td>$0.48$</td>
<td>$0.45$</td>
</tr>
<tr>
<td><strong>Partial Correlations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faces</td>
<td>$-0.90$</td>
<td>$-0.28$</td>
<td>$-0.15$</td>
<td>$-0.73$</td>
</tr>
<tr>
<td>Vehicles</td>
<td>$0.83$</td>
<td>$0.37$</td>
<td>$0.43$</td>
<td>$0.33$</td>
</tr>
</tbody>
</table>

Zero-order Pearson correlation coefficients are shown for face and vehicle categories correlated with total regional thickness and cortical depths operationalized by the fourth-order polynomial and its second-order derivative. Partial-order Pearson correlation coefficients are given for each behavioral category regressing out the other two (including animals), correlated with total regional and laminar cortical depths. Significant correlations are **bolded**.

**Figure 3.** The relation between GM thickness of rFFA2 or its laminar subdivisions, and behavioral performance. Behavioral performance on the $x$ axes and CT and depth measurements on the $y$ axes show residualized values (see text for actual values in mm). The scatter plots on the left (with face icons) show partial correlations with FR controlling for performance with vehicles, whereas scatter plots on the right (with car icons) show partial correlations with vehicles controlling for FR. (A) The brain is tilted as a result of slices being aligned perpendicular to the rFFA2. Shown is one slice for one participant showing the average of six susceptibility-weighted scan acquisitions. The dashed box is centered on the functionally defined rFFA2. Below the whole-brain slice are scatter plots depicting the relation between total regional CT of rFFA2 and behavioral performance with faces (left) and vehicles (right). Error bars represent the 95% confidence intervals. (B) The central inset shows a magnified view from the box in (A), centered on rFFA2 for one slice in one participant, with the three laminar subdivisions perceivable to the naked eye. Scatter plots represent partial correlations between the thickness of deep, middle, or superficial MR layers as a function of FR and VR, controlling for performance with the other category. Error bars show the 95% confidence intervals.
ps ≤ .001) and together accounted for 60% of the variance on adjusted $R^2$.

As in prior work (McGugin et al., 2016), the correlation between overall CT and recognition of animals, which may not be acquired especially early or late, was in the same direction as FR, although not significantly so ($t = −.12, p = .34$). Animal recognition also did not predict thickness of MR layers (superficial: $r = .15$; middle: $r = .26$; deep: $r = .26$). Our results cannot be explained by variability in vascular volume in FFA2 because such differences were not related to total CT ($t = −.31$), or behavioral performance (face, $r = .11$; vehicle, $r = −.09$, animal, $r = .21$), despite being reliable (ICC = .75).

**Association between Behavior Measures and Thickness of rFFA2 MR Layers**

**Correlations**

In accord with predictions of the differential early development model, FR was strongly negatively correlated with the thickness of deep MR layers of rFFA2 ($r = −.76, p = .002$; Figure 3B; Table 2). This relation is evident when examining participant-specific brain data listed as a function of FR performance (Figure 4). There were no significant correlations between FR and thickness of superficial or middle MR layers (Figure 3B; Table 2). In contrast, correlations between VR and thickness of all three MR layers were consistently positive in sign ($rs$ ranged from .44 to .48, two-tailed $ps$ ranging from .13 to .10), although not statistically significant. For the deep MR layer, the correlations of thickness with VR significantly differed from that with VR ($t = 3.69, p < .005$). These effects were replicated when bootstrapping was used.

**Regression Models Predicting Behavioral Measures from MR Layer-specific CT**

Although correlations revealed bivariate relations among pairs of variables, they do not address the degree to which the combination of MR layer-specific CT measures predict individual differences in behavior. To address this question, we conducted multiple regression analyses in which CT in the superficial, middle, and deep MR layers predicted FR and VR scores. According to adjusted $R^2$, the combination of superficial, middle, and deep MR layer thickness accounted for 65% of the variance in FR (overall regression, $F(3, 9) = 8.29, p = .006$, unadjusted $R^2 = .73$). Both standard and bootstrapped confidence intervals indicated that the thickness of the deep layer was the only statistically significant predictor of FR. We also computed semipartial correlations indicating the correlation between FR and each CT measure with the latter adjusted for the other CT measures. Only the semi-partial correlation for the deep MR layer ($r = .72, p = .0125$) was significantly greater than 0 (superficial $r = −.29, p = .40$; middle $r = −.25, p = .45$). Squared semipartial correlations quantifying the proportion of variance uniquely contributed by each predictor indicated that the thickness of the deep MR layer accounted for slightly over half the total variance of FR (deep $sr^2 = .52$, superficial and middle $sr^2 = .08$ and .06, respectively).

The multiple regression analyses with VR as the dependent variable had a different pattern of results. In this case, the overall regression model was significant, $F(3, 9) = 3.98, p < .05$, although none of the regression coefficients for the three predictors were significant ($ps$ ranged from .07 to .12). Although the latter result could be due to the relatively small sample size, this result also suggests that it was the shared variance among the three predictors rather than the unique variance of each predictor that accounted for effects on VR.

**Regression Models Predicting CT of MR Layers from Behavior**

Our most extensive analyses used a multiple regression framework to assess alternative models of the effects of FR and VR on CT measures. We emphasized these analyses because they allowed us to test alternative hypotheses about the complete pattern of associations between the two behavioral measures and thickness of the three MR layers.

Our regression models used a repeated-measures framework, with the three MR layer-specific CT measures per person serving as the repeated dependent measure. Predictors were layer, FR, and VR. We created two new variables that decomposed the overall effect of the layer factor into two orthogonal (i.e., uncorrelated) contrasts. As noted above, we hypothesized that the association between FR and thickness would be more strongly negative in the deep MR layer relative to the superficial and middle MR layers. To capture this difference, we gave one contrast variable (denoted $L_{DvsSM}$) a code of −1 for observations from the deep MR layer and codes of 0.5 for observations from the superficial and middle MR layers. The second contrast variable for layer (denoted $L_{SmM}$) was orthogonal to the first and assigned a code of 0 to the deep layer, 1 to the superficial layer, and −1 to the middle layer. When both contrast variables are included in a model as a set, they estimate the overall (2 df) main effect of MR layer.

We compared six different regression models that imposed different restrictions on the association between measures of the CT of MR layers (hereafter “layers” for short in the model descriptions) and FR and VR.

**Model 0 ($M_0$): Baseline model.** In this model, the only predictor of CT levels was the overall main effect of layer,

$$\text{Model 0 : } \text{CT}_{ij} = B_0 + B_1 L_{DvsSM} + B_2 L_{SmM} + \varepsilon_{ij}$$

where $i$ denotes participant (1 through 13), $j$ denotes layer (1 through 3), and $\varepsilon$ denotes the prediction error for the CT measure in the $j$th layer for the $i$th participant.
Figure 4. Plots of the fourth-order polynomial fit to the average rFFA2 intensity curve of the center-most slice for each participant. Participants are rank-ordered from highest (S1) to lowest based on FR performance, regressing out performance with vehicles. Inflection points marking the transitions from deep MR layers (red shading) to middle MR layers (no shading) to superficial MR layers (gray shading) are marked with red dots. Axis labels: WM, CSF. For every other participant, the centermost slice of rFFA2 is shown to the left. The spread of face-selective activation along the fusiform gyrus is outlined in gray. Labels: OTS = occipital temporal sulcus; CoS = collateral sulcus; MFS = middle fusiform sulcus. The rFFA2 most often fell on the lateral aspect of the fusiform gyrus, between MFS and OTS, corresponding to cytoarchitectonic FG4.
We used M3 simply as a baseline comparison to verify that FR and VR measures, introduced in subsequent models, were significantly associated with CT.

Model 1 (M1): Equal coefficients across layers. The next two models reflected our two primary alternative hypotheses. M1 predicted CT in a given layer from the two contrast variables for layer, FR, and VR,

\[CT_{ij} = B_0 + B_1L_{DvsSM} + B_2L_{SvM} + B_3FR_i + B_4VR_i + \epsilon_{ij}\]

M1 is a “main effects only” model with no interaction terms. As a result it specifies that the regression coefficient denoting the effect of FR on CT (B3) is equal (i.e., invariant) across all three layers and the coefficient for the effect of VR on CT (B4) is also equal across layers (see Figure 1). This model is the most parsimonious extension of our previous finding using overall CT (McGugin et al., 2016).

Model 2 (M2): FR differentially predicts deep versus superficial/middle MR layer CT. M2 tested the hypothesis that (1) the association between FR and CT is more pronounced (i.e., more negative) in the deep layer than the middle and superficial layers, but (2) layer does not moderate the relation between VR and CT measures. To embody this specification, an interaction (i.e., product) term between FR and the deep versus superficial/middle contrast variable (L_{DvsSM}) was added to M1,

\[CT_{ij} = B_0 + B_1L_{DvsSM} + B_2L_{SvM} + B_3FR_i + B_4VR_i + B_5(FR_i \times L_{DvsSM}) + \epsilon_{ij}\]

In M2, three unique regression coefficients are sufficient to account for the linkages between FR and VR and CT across layers: the coefficient for the regression of CT on FR within the deep layer (equal to B3 – B5), the coefficient for the regression of CT on FR within the middle and superficial layers (equal to B3 + B5/2), and the single coefficient for the regression of CT on VR within each of the three layers (B4). Thus, this model specifies a constant effect of VR across all three layers and the hypothesized variable effect for FR.

The remaining models are designed to assess patterns of association other than those directly hypothesized in Models 1 and 2.

Model 3 (M3): Extended M2 by adding an interaction term between the second layer contrast variable (L_{SvM}) and FR. This addition relaxes the assumption that the regression coefficients for FR in the middle and superficial layers were equal to one another. Thus, M3 goes beyond M2 by allowing all three regression coefficients for FR to differ from one another. Like M2, however, it restricts all three regression coefficients for VR to be equal.

Model 4 (M4): Added the full VR \times Layer interaction to M1. This model specifies that the regression coefficients involving FR are all the same but allows for unique regression coefficients for VR in each layer.

Model 5 (M5): Extended M2 and M4 by adding in both the overall FR \times Layer interaction and the overall VR \times Layer interaction. This model allows for six unique regression coefficients (FR/VR \times Superficial/Middle/Deep) to capture the association between FR and VR and CT measures. Support for this model would imply that layer moderates the relation between both FR and VR and CT.

We conducted analyses using the gls function (Galecki & Burzykowski, 2013) in the nlme package (Pinheiro et al., 2018) in R (R Core Team, 2018). For all models, we accounted for the dependence among the three observations for each participant by imposing an unstructured covariance matrix among the residuals. Maximum likelihood estimation was used (e.g., Galecki & Burzykowski, 2013). We evaluated the relative fit of the models using several indices. When models were nested, we used likelihood ratio tests to compare models. We also used three information indices to compare all models: the Akaike Information Criterion (Akaike, 1973), the small-sample corrected Akaike Information Criterion (Sugiura, 1978), and the Bayesian Information Criterion (Schwarz, 1978). These indices reward model fit but penalize for model complexity and thus are designed to prevent overfitting. Lower values indicate better fit. We also report −2 \times the log likelihood (−2LL) value of each model computed at the maximum likelihood estimates. −2LL values are used both in the computation of likelihood ratio tests (for which the difference in the −2LL values of two models are distributed as chi-square variables with degrees of freedom equal to the difference in the number of parameters estimated by each model) and of information indices.

Table 3 shows the values of −2LL and the information indices for the six models. The values for M0 very clearly fit worse than all other models (e.g., all LR tests comparing other models to M0 p < .005), thus indicating that FR and VR contributed significantly to the prediction of CT. Most importantly, Table 4 indicates that M2 (FR Differentially Predicts Deep versus Superficial/Middle CT) is clearly the best-fitting model. It has the lowest values for all three information indices. Likelihood ratio tests comparing M2 to other models yielded a pattern consistent with the information indices. M2 fit significantly better than M1 (\(\chi^2 = 5.94, p = .02\)). Although M2 is a restricted version of M3 and M4, neither fit any better than M2 (M2 vs. M3: \(\chi^2 = .01, p > .90\); M2 vs. M4: \(\chi^2 = .20, p > .97\)). These results indicate that M2 is the model that best combines fit to the data with parsimony. Table 4 shows the unstandardized and standardized regression coefficients for the regression of CT on FR and VR that are implied by M2. FR has a stronger (negative) effect on CT in the deep MR layer relative to the superficial
and middle MR layers. This difference is statistically significant, $t(33) = 2.53, p < .02$.

Although the fit indices discussed so far clearly indicate that M2 is the best-fitting model, they do not precisely quantify the strength of the evidence for M2 relative to the other hypotheses. For this reason, we additionally computed Bayes factors (BFs; Kass & Raftery, 1995; Jeffreys, 1961) using the BayesFactor package in R (Morey et al., 2018). We specified the same six regression models. Instead of correlated residuals, we specified a random effect for person. For all models, we used the Jeffreys–Zellner–Snow mixture of $g$-priors specification introduced by Liang, Paulo, Molina, Clyde, and Berger (2008; see also Rouder & Morey, 2012). Because BFs can be sensitive to prior distributions, we used three alternative specifications of the scale parameter of the inverse gamma distribution that, in turn, is used to specify the variances of the prior distribution for the regression coefficients. All three specifications produced identical conclusions. Below, we report the results for the default $g$-prior, which yielded the highest evidence favoring M2.

An examination of the BFs yielded by the 15 pairings of models indicated that M2 was clearly the best-fitting model (i.e., for each of its five comparisons, M2 had the best fit). The final column in Table 3 displays the BFs comparing M2 to each of the other models. Perhaps the most important index here is the BF of 38.85 for the M2/M1 comparison. This value indicates that the evidence in the data for M2 is over 30 times as strong as the evidence for M1. All other comparisons yielded what Jeffreys’ termed “strong” support for M2 except that involving M4. This is not surprising given that both models are highly similar. Both allow for a difference in the association between FR and CT between the deep and the superficial layers, and both specify that the regression coefficients for VR are the same across layers.

Table 3. Fit Indices for Regression Models Predicting CT

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>$\sim 2LL$</th>
<th>AIC</th>
<th>AICc</th>
<th>BIC</th>
<th>M2/M# BF</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0</td>
<td>Layer only</td>
<td>43.97</td>
<td>61.97</td>
<td>68.18</td>
<td>76.95</td>
<td>161.04</td>
</tr>
<tr>
<td>M1</td>
<td>$M_0 + \text{Main effects for FR and VR}$</td>
<td>13.51</td>
<td>35.51</td>
<td>45.29</td>
<td>53.81</td>
<td>38.85</td>
</tr>
<tr>
<td>M2</td>
<td>$M_1 + L_{\text{DVS}}$</td>
<td>7.56</td>
<td><strong>31.56</strong></td>
<td><strong>43.56</strong></td>
<td><strong>51.53</strong></td>
<td>—</td>
</tr>
<tr>
<td>M3</td>
<td>$M_1 + \text{Layer × FR Interaction}$</td>
<td>7.54</td>
<td>33.56</td>
<td>48.12</td>
<td>55.18</td>
<td>3.07</td>
</tr>
<tr>
<td>M4</td>
<td>$M_1 + \text{Layer × VR interaction}$</td>
<td>12.52</td>
<td>38.52</td>
<td>53.09</td>
<td>60.15</td>
<td>139.52</td>
</tr>
<tr>
<td>M5</td>
<td>$M_1 + \text{Layer × FR + Layer × VR interactions}$</td>
<td>7.34</td>
<td>37.34</td>
<td>58.21</td>
<td>62.29</td>
<td>19.90</td>
</tr>
</tbody>
</table>

$\sim 2LL = -2 \times \log \text{likelihood of the model.}$ This index is used in the computation of likelihood ratio tests comparing nested models and information indices. As lower values of information indices indicate better fit, all favor $M_2$ (in bold). BFs compare $M_2$ to each of the other models. All BFs indicate greater support for $M_2$ than all other models. AIC = Akaike Information Criterion; AICc = AIC with small-sample correction; BIC = Bayesian Information Criterion.

Table 4. Unstandardized and Standardized Regression Coefficients from $M_2$

<table>
<thead>
<tr>
<th>Layer</th>
<th>FR</th>
<th>VR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superficial</td>
<td>$-.07 (-.10)$</td>
<td>$.18^* (.27)$</td>
</tr>
<tr>
<td>Middle</td>
<td>$-.07 (-.10)$</td>
<td>$.18^* (.27)$</td>
</tr>
<tr>
<td>Deep</td>
<td>$-.62^* (-.57)$</td>
<td>$.18^* (.27)$</td>
</tr>
</tbody>
</table>

$M_2$ is the best-fitting regression model that predicts that the FR is selectively associated with CT in the deep layer but that VR has the same association with CT in each layer. Unstandardized regression coefficients are on top and standardized regression coefficients are in parentheses.

**DISCUSSION**

We used susceptibility-weighted MRI in a 7T magnet to measure the thickness of deep, middle and superficial MR layers, distinctions that cannot be resolved with conventional MRI approaches at 3T or below. Although new developments in MRI have allowed some visualization of intracortical layers, validation of their functional relevance has been reserved to primary visual cortex (Trampel, ...
Ott, & Turner, 2011). We were constrained to a single ROI because of a limited field of view and the fact that accurate measurement of laminar CT with nonisotropic voxels depends upon careful alignment of slices perpendicular to the cortex of a targeted brain area. We selected the rFFA2 based on earlier robust functional expertise effects for nonface objects (McGugin, Van Gulick, Tamber-Rosenaau, Ross, & Gauthier, 2015).

Here, we replicate and extend the finding that VR is positively correlated with CT of FFA (McGugin et al., 2016), whereas FR is negatively correlated with FFA thickness (McGugin et al., 2016; Bi, Chen, Zhou, He, & Fang, 2014). This surprising pattern of correlations led us to postulate different mechanisms of plasticity behind these effects, with the mechanism relevant to FR operating earlier and selectively influencing the deep layers of rFFA2.

To measure the thickness of MR layers, we applied an automated algorithm to define borders between layers using signal intensity shifts along perpendicular traces through rFFA2 (Methods and Figure 2). We postulate that the three most prominent intensity differences are related to observed laminar divisions in the fusiform (Lorenz et al., 2017), corresponding to the densely-packed deep/infragranular layers (layers V–VI), middle/internal granular (layer IV), and superficial/supragranular (layers I–III). We successfully visualized laminar structure in rFFA2 and our measurements of the thickness of MR layers were reliable.

Although we compared several different models of the data, our primary focus was on two specific alternatives (see Figure 1). The first (M1) specified that the pattern of correlations observed in total CT (negative for faces and positive for vehicles) is constant across all laminar subdivisions. The second (M2) specified that the association between individual differences in FR and the thickness of the deep MR layer was stronger (i.e., more negative) than the association between FR and thickness of the superficial and middle MR layers. Like M1, M2 specified that the association between VR and CT was invariant across layer. Regression analyses consistently favored M2 relative to M1 and to the other models that we specified, across multiple indices (likelihood ratio tests, three information indices, and BFs). Comparisons of regression coefficients yielded by M2 indicated that the association between FR and the deep MR layer was significantly stronger than the association between FR and the superficial and middle MR layers. The thickness of the deep MR layer accounted for a much higher proportion of variance in FR than the thickness of the other two MR layers. Indeed, when considered in isolation, CT of the deep MR layer accounted for 58% of the variance in FR, which is 84% of the variance accounted for by all the layers together (using total CT). Thus, the strong association between total CT and FR is largely, if not exclusively, due to the deep MR layer. In contrast, for VR there was no evidence that the relation was moderated by layer, and the clearly best-fitting regression model (M2) constrained the regression coefficients for VR to be equal across layers.

The sample size of this study (n = 13) was small. Accordingly, we focused on regression models that specified the overall pattern of association across layers and behavioral measures, to increase power and precision (e.g., see Hoijtink, 2011). In the present case, for example, although the individual correlations between VR and CT in the superficial, middle, and deep layers were not significantly greater than 0, our regression models yielded significant effects for VR on CT when these were constrained to be equal across layers. More generally, that our analyses indicated the clear superiority of M2 even given the relatively small sample size is an indication of its strength as a representation of the underlying pattern of associations.

These results are among the most structurally precise correlates of behavioral ability to date and an important validation of the functional relevance of our in vivo laminar measurements. Future development of these methods to a larger field of view should allow exploration of how these fine structural effects relate to correlations of FR ability on a much more distributed scale (Elbich & Scherf, 2017). Our results also represent perhaps the most striking difference to date between FR and non-FR. FR has often been deemed special relative to the recognition of other categories (Farah, Wilson, Drain, & Tanaka, 1998), but many behavioral hallmarks of face expertise, such as the inversion effect or holistic processing, are observed with non-face objects like cars in expert participants (Bukach, Phillips, & Gauthier, 2010; Curby, Glazek, & Gauthier, 2009). Face and car recognition also show similar degrees of heritability (Shakeshaft & Plomin, 2015) and similar correlations with domain-general object recognition ability (Richler, Wilner, & Gauthier, 2017). In fMRI studies, car recognition ability correlates positively with selective neural responses to images of cars in FFA (e.g., McGugin, Gatenby, et al., 2012), whereas FR ability also correlates positively with selective responses to faces in the same area (McGugin, Ryan, Tamber-Rosenaau, & Gauthier, 2018). Despite these similarities, our results suggest that behaviors supported by common circuitry in a brain region may have developed at different times, with variability in structure as a footprint of the different acquisition history.

Our predictions were inspired by the idea that FR ability begins earlier than VR and that a skill acquired during this period would be influenced by the pruning of connections from IT to limbic areas, though this is not the only possibility. We do not know whether the relevant period of influence on FFA’s structure occurs before about 5 years of age (by that age, children show many signs of the expertise adults display with faces; Jeffery et al., 2011), when pruning may be more important, or in later childhood and adolescence, when myelination may dominate developmental changes (Natu et al., 2018). An alternative explanation is that greater myelination near FFA during childhood and adolescence is associated with better FR. Recent work has questioned the prevalent theory that pruning is the main force behind cortical thinning during
development, with more evidence for myelination than for pruning after 5 years of age in the fusiform gyrus (Natu et al., 2018). Also unknown are the specific mechanisms of plasticity that would support structural effects of visual learning in adults, leading us to assume the null hypothesis of no laminar specificity of the positive correlation with CT observed for vehicles. Importantly, the mechanism driving the largest amount of average change in CT in an area may not be the same mechanism driving individual differences in CT.

It is unknown whether negative correlations between CT and performance will prove to be specific to deep MR layers for other kinds of abilities and other brain regions. It is also unknown which of the two patterns (positive correlation with total regional CT or negative correlation with thickness of deep MR layers) is more representative of the majority of object categories. Here, as in prior work at 3T (McGugin et al., 2016), there was no significant correlation with animal recognition but the trend was more similar to effects found with faces than vehicles (a negative correlation for thickness of deep MR layers). Future studies should explore how laminar CT in FFA relates to a more extensive set of visual abilities and how this relation is affected by sex. Indeed, as of now, the crossover interaction between CT and visual recognition abilities has only been described in men, yet there are sex differences in the recognition of living and nonliving objects (McGugin, Richler, et al., 2012), and sex also moderates CT effects (Plessen, Hugdahl, Bansal, Hao, & Peterson, 2014).

By the time CT is measured in adults, effects in any given brain area will reflect a complex history of changes due to both experience- and non-experience-dependent developmental factors. This complicates interpreting relationships between behavior and local brain structure, including how structure underlies various neurological disorders. For example, both abnormal thinning of fusiform cortex over development (Raznahan et al., 2010) and lack of developmental improvements in FR (O’Hearn, Schroer, Minshew, & Luna, 2010) have been reported in autism spectrum disorders. The study of the laminar specificity of structural correlates of behavioral abilities may provide a useful window into these complex dynamics.

Acknowledgments

We thank Carla Shatz for suggesting that different ages of acquisition should translate into differential effects on cortical layers. We thank Magen Speegle and Susan Benear for manual tracing on the gray matter–white matter boundaries; Anita Disney, Jesse Gomez, Thilo Womelsdorf, and Wolfe Zinke for comments; Maarten Versluis and Ed Mojahe d for sharing pulse sequence designs that were used as a basis for those developed here. R. W. M. and I. G. conceived the study. R. W. M. and A. T. N. collected the primary data. A. T. N. developed the image acquisition techniques and average susceptibility-weighted image reconstruction techniques. R. W. M., A. T. N., and B. T. R. developed the methods for laminar analysis of susceptibility-weighted images. A. T. conducted reliability, bootstrapping, and Bayesian analyses. All authors coordinated second-level analyses. R. W. M., A. T. N., and I. G. wrote the article with input from the other authors. This work was supported by the NSF (SBE-0524013 and SMA-1640681 and NSF award BCS-1848096), the Vanderbilt Vision Research Center (P50-EY008126), and the Vanderbilt University Medical Center (1 S10RR023047 01).

Reprint requests should be sent to Rankin W. McGugin, Department of Psychology, Vanderbilt University, Wilson Hall, PMbx 407817, 2301 Vanderbilt Place, Nashville, TN 37240-7817, or via e-mail: Rankin.McGugin@gmail.com.

REFERENCES


