Structural Differences in Adult Orbital and Ventromedial Prefrontal Cortex Predicted by Infant Temperament at 4 Months of Age

Carl E. Schwartz, MD; Pratap S. Kunwar, MS; Douglas N. Greve, PhD; Lyndsey R. Moran, BA; Jane C. Viner, BA; Jennifer M. Covino, BS, MPA; Jerome Kagan, PhD; S. Evelyn Stewart, MD; Nancy C. Snidman, PhD; Mark G. Vangel, PhD; Stuart R. Wallace, BS

Context: The term temperament refers to a biologically based predilection for a distinctive pattern of emotions, cognitions, and behaviors first observed in infancy or early childhood. High-reactive infants are characterized at age 4 months by vigorous motor activity and crying in response to unfamiliar visual, auditory, and olfactory stimuli, whereas low-reactive infants show low motor activity and low vocal distress to the same stimuli. High-reactive infants are biased to become behaviorally inhibited in the second year of life, defined by timidity with unfamiliar people, objects, and situations. In contrast, low-reactive infants are biased to develop into uninhibited children who spontaneously approach novel situations.

Objective: To examine whether differences in the structure of the ventromedial or orbitofrontal cerebral cortex at age 18 years are associated with high or low reactivity at 4 months of age.

Design: Structural magnetic resonance imaging in a cohort of 18-year-olds enrolled in a longitudinal study. Temperament was determined at 4 months of age by direct observation in the laboratory.

Setting: Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital.

Participants: Seventy-six subjects who were high-reactive or low-reactive infants at 4 months of age.

Main Outcome Measure: Cortical thickness.

Results: Adults with a low-reactive infant temperament, compared with those categorized as high reactive, showed greater thickness in the left orbitofrontal cortex. Subjects categorized as high reactive in infancy, compared with those previously categorized as low reactive, showed greater thickness in the right ventromedial prefrontal cortex.

Conclusions: To our knowledge, this is the first demonstration that temperamental differences measured at 4 months of age have implications for the architecture of human cerebral cortex lasting into adulthood. Understanding the developmental mechanisms that shape these differences may offer new ways to understand mood and anxiety disorders as well as the formation of adult personality.

Arch Gen Psychiatry. 2010;67(1):78-84
ferences in activity in circuits that project from the amygdala to the sympathetic chain and suggest that the complex behavioral and physiological profiles of these 2 temperaments might reflect differential excitability of the amygdala.

The 2 infant temperaments are associated with distinct psychological features in adolescents. Fifteen-year-olds who had been high-reactive infants showed subdued social behavior, right frontal electroencephalogram activation, greater sympathetic over parasympathetic tone, and a shallower habitation of the event-related potential at 400 milliseconds to discrepant visual events. Adolescents who had been low-reactive infants showed spontaneous social behavior, left frontal activation, vagal dominance, and a steeper habitation of the event-related potential to discrepant visual events. In addition, more high-reactive than low-reactive adolescents reported serious worry over encounters with unfamiliar situations and more frequent melancholic moods. More low-reactive than high-reactive adolescents reported worrying only over realistic events, such as school grades and athletic performance, and reported happier moods. Independent prospective studies from several laboratories have demonstrated that an inhibited temperament is a risk factor for the development of anxiety disorder in childhood and adolescence, particularly social anxiety disorder. Social anxiety disorder during adolescence in turn is an important predictor of subsequent depressive disorders and social phobia in young adults.

A previous functional magnetic resonance imaging (MRI) study from our laboratory supported the hypothesis that the differences in physiology and behavior between inhibited and uninhibited temperaments might indeed reflect differential amygdalar reactivity to novelty. There is rich bidirectional connectivity between the ventral prefrontal cortex and the amygdala. The ventral prefrontal cortex, including the orbitofrontal cortex (OFC), plays a pivotal role in emotional regulation, reward processing, and the ability to inhibit behaviors. We therefore wondered if differences in the thickness of the ventral prefrontal cortex in adults would differentiate high-reactive from low-reactive infants. Using high-resolution structural MRI, we tested this hypothesis in 76 subjects who were enrolled in an 18-year longitudinal study and had been characterized as high-reactive (n = 34) or low-reactive (n = 42) infants at 4 months of age (Table 1). Handedness, measured with the Edinburgh Inventory, did not differ between the 2 temperament groups. Twenty-two of the high-reactive infants in this sample were also categorized as highly fearful (ie, inhibited) children in the second year of life, whereas just 5 children were classified as having low fear (uninhibited). Similarly, 26 of the low-reactive infants were also categorized as having low fear in the second year, whereas just 3 were classified as being highly fearful.

### METHODS

#### INFANT ASSESSMENT AND CATEGORIZATION

The details of the standard 45-minute battery are described elsewhere in detail. Initially, the mother looked down at her infant smiling, but not talking, for 1 minute. The parent then went to a chair behind the infant to be outside the child’s field of vision. The examiner then placed a speaker baffle to the right of the infant and turned on a tape recording that played 8 short sentences read by female voices. The speaker baffle was removed and the examiner, standing in back of the infant, presented a set of mobiles composed of 1, 3, or 7 colorful toys that moved back and forth in front of the infant’s face for 9 second trials. The examiner then dipped a cotton swab into very dilute butyl alcohol and presented it close to the infant’s nostrils for 8 trials (the first and last trials were water rather than alcohol). The speaker baffle was replaced and the infant heard a female voice speaking 3 nonsense syllables (ma, pa, ga) at 3 different loudness levels. The examiner then popped a balloon in back of the infant; most were unperturbed by this event. Finally, the mother returned to gaze at her infant for the final minute. The decision to define discrete groups based on the combination of motor activity and crying, rather than a continuum of reactivity, was supported by a taxonomic analysis of the 4-month data that implied that the combination of the 2 variables fit a categorical model better than a continuous one.

### NEUROIMAGING

Each subject underwent two 3-dimensional MPRAGE structural scans on a 3-T Siemens (Malvern, Pennsylvania) TrioTim scanner (128 sagittal slices; 1.3 × 1.3 × 1 mm; repetition time = 2530 milliseconds; echo time = 3.39 milliseconds; flip angle, 7; bandwidth, 190 Hz/px). The two 3-dimensional MPRAGE structural scans from each subject were averaged, after motion correction, to create a single high–signal-to-noise average volume. This volume was analyzed using FreeSurfer (www.nmr.mgh.harvard.edu/...both to create anatomical surface models and perform statistical analyses. The details of these methods have been reported elsewhere. The average volume for each subject was used to create a finite-element surface mesh model of the cortical surface, both at the gray matter–white matter junction and pial surface. The gray matter–white matter boundary and pial surfaces of each subject were carefully examined and edited to ensure fidelity to each individual’s anatomy. Each element in this model is called a "vertex." For each subject, thickness measures across the cortex were computed by finding the point on the gray matter–white matter boundary surface that was closest to a given point on the estimated pial surface (and vice versa) and averaging between these 2 values. The accuracy of the thickness measures derived from this technique has been validated by direct comparisons with manual measures on postmortem brains.

To map each subject to a common space, the surface representing the gray matter–white matter border was registered to an average cortical surface atlas using a nonlinear procedure that optimally aligned sulcal and gyral features across subjects. Cortical parcels were drawn on the anatomical atlas; parcelations were mapped back onto each individual subject’s surface by applying the subject-atlas registration described earlier. For the vertex-by-vertex cluster analysis, the thickness maps for all subjects in both groups were converted to the common atlas.
The maximal thickness difference between the 2 groups defined by infant temperament occurs at the vertex with Talairach coordinates \(-24, 35, \text{ and } -10\), and is indicated with a light blue point. This vertex can be seen in the wall of the transverse orbital sulcus in the noninflated (pial) view in part A but is easier to appreciate in the inflated image of the orbitofrontal surface in part B.

The maximal thickness difference between the 2 groups in the right hemisphere occurs at the vertex with Talairach coordinates \(6, 46, \text{ and } -16\), and is indicated with a light blue point in the medial view of the right hemisphere.

The data were smoothed on the surface using an iterative nearest-neighbor averaging procedure (74 iterations were applied, equivalent to applying a 2-dimensional Gaussian smoothing kernel along the cortical surface with a full-width half-maximum of approximately 10 mm). A general linear model was used to test for cortical thickness differences between the 2 temperament groups and the 2 sexes and for any interaction between these 2 factors at each vertex. To correct for multiple comparisons, spatial clusters of thickness differences were defined as contiguous patches of vertices with \(P\) values less than .05 (2-tailed). The \(P\) values for these clusters were determined by Monte Carlo simulation (10,000 iterations). Only clusters that survived this correction with \(P\) values less than .05 (2-tailed) were deemed significant. For \(P = .05\), the cluster size threshold in the combined search area of the ventral prefrontal cortex (consisting of the frontal pole, ventromedial and ventrolateral prefrontal cortex including the OFC, and the pars orbitalis) was 168 mm\(^2\). We also performed a vertex-wise whole-brain analysis to examine whether there were any additional areas of thickness differences between the 2 temperament groups that survived correction for multiple comparisons at the whole-brain level. In addition, the posterior visual cortex (cuneus, pericalcarine, and lingual gyrus) was selected a priori as a comparison region that we predicted would not show a significant difference in cortical thickness between groups. Finally, we examined whether there were spatially diffuse thickness differences between the 2 temperamental groups in the ventral prefrontal cortex in addition to the confluent clusters of thickness differences. The mean thickness of the ventral prefrontal cortex (frontal pole, ventromedial and ventrolateral prefrontal cortex including the OFC, and the pars orbitalis), but clipping out the territory of the clusters in Figure 1, was computed for each individual. This computed thickness was the dependent variable in a general linear model, with temperament type (2 levels) and sex as between-subject factors. To exclude a difference in the goodness of fit to the common atlas space as a potential source of bias in the comparison of the 2 temperament groups, we compared the curvature index in the 2 groups, since this index is used as the basis of surface registration; no difference was found. The spherical coordinate space in which each subject’s cortex is registered to the atlas after inflation in FreeSurfer is particularly well suited to handling variability in sulcal and gyral anatomy among individuals. This has been cited as an advantage of FreeSurfer compared with other approaches to MRI analysis. Results were visualized on a group brain generated by the actual subjects in the study rather than an average atlas brain, thereby reflecting more accurately any distinctive anatomical variation. This enabled more accurate description of the spatial location of clusters with respect to gyral and sulcal features. Data analysts were blind during image processing to subjects’ identity and temperamental type in infancy.

Subjects with a low-reactive temperament at age 4 months had a thicker cortex in a region of the left OFC compared with those with a high-reactive temperament, whereas high-reactive subjects had a thicker cortex than low-reactive subjects in a region of the right ventromedial prefrontal cortex (Figure 1 and Figure 2). There was no difference in cortical thickness between the sexes, nor any interaction between temperament and sex in either of these regions (Table 2).

Figure 1 illustrates the 225-mm\(^2\) region of the left OFC that was thicker in the 18-year-olds who had a low-reactive temperament, compared with those who were high-reactive infants. The point of maximal thickness difference between the 2 temperamental groups, marked with a bright blue spot, lies in the transverse orbital sulcus. The cluster extends into the anterior-lateral portion of the posterior orbital gyrus, the most extreme lateral aspect of the medial orbital gyrus, the most extreme medial aspect of the lateral orbital gyrus and pars orbitalis, and the most posterior aspect of the ant-
The 2 major contemporary maps of the human OFC are by Petrides and Mackey and Petrides and Pandya and by Price and Ongür et al. In the Petrides and colleagues map, the location of the left orbitofrontal cluster corresponded primarily to area 13, bounded by a transitional zone between 13 and 47/12 laterally, areas 13 and 11 anteriorly, and the junction of 13 and 14 medially. In the more fine-grained schema of Price and Ongür et al, the cluster corresponded to the cortex in areas 47/12m, 13l, and 11l and was bounded by the transitions between 47/12m and 47/12 laterally, 47/12m and 11l anteriorly, and the junction of 13m and 13l medially.

The right ventromedial cluster, which was thicker in the high-reactive subjects, lay on the medial wall of the cerebral hemisphere and corresponded in the Petrides and colleagues map to the limbic cortex within areas 14 in the inferior/posterior part of the cluster and area 32 in the superior/anterior part of the cluster. In the Price and Ongür et al map, the most posterior aspect of the cluster corresponded to area 14r; the cluster extended into the most posterior aspect of area 11m and the most inferior/anterior corner of area 10m, before reaching area 10r at the most superior and anterior aspect.

Because social anxiety disorder in adolescence has been linked to an inhibited temperament, we asked whether our results might be due to a confounding with the use of medication, social anxiety disorder, or major depressive disorder. The results indicated that the thickness differences between the temperament groups were not associated with any of these factors (Table 4).

These data suggest that regional differences in the thickness of adult OFC and ventromedial prefrontal cerebral cortex are predicted by temperamental differences observed in infancy. This cluster bridges several anatomical regions as defined by surface gyral and sulcal anatomy. Figure 2 illustrates the 169-mm² region of the right ventromedial prefrontal cortex that was thicker in the 18-year-olds who had a high-reactive temperament in infancy, compared with those subjects who were low-reactive infants. This cluster is located on the medial wall of the gyrus rectus of the ventromedial prefrontal cortex. The cluster extends diagonally across the medial wall of the rectus gyrus angled upwards from its most inferior/posterior territory to the most superior aspect of the cluster that is more anterior. The most superior aspect extends to include cortex lying within the superior rostral sulcus, which defines the most superior aspect of the cluster that is more anterior.

Table 3 shows the range of Talairach coordinates that occurred in the regions illustrated in Figure 1 and Figure 2 and the Talairach coordinates of the vertex at which the thickness difference between the temperament groups was greatest.

Table 3. Range and Maxima of Talairach Coordinates of Regional Cortical Thickness Differences Between Temperaments in Figure 1 and Figure 2

<table>
<thead>
<tr>
<th>Right Ventromedial PFC (Thicker in High-Reactive Subjects Than Low-Reactive Subjects)</th>
<th>Left OFC (Thicker in Low-Reactive Subjects Than High-Reactive Subjects)</th>
</tr>
</thead>
<tbody>
<tr>
<td>x-Axis Med to Lat</td>
<td>y-Axis Post to Ant</td>
</tr>
<tr>
<td>Range</td>
<td>4 to 11</td>
</tr>
<tr>
<td>Maximum</td>
<td>6</td>
</tr>
</tbody>
</table>

Abbreviations: Ant, anterior; Inf, inferior; Lat, lateral; Med, medial; OFC, orbitofrontal cortex; PFC, prefrontal cortex; Post, posterior; Sup, superior.

Table 2. Cortical Thickness by Temperament

<table>
<thead>
<tr>
<th>Right Ventromedial PFC, mm, Mean (SD)</th>
<th>Left OFC, mm, Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-Reactive Subjects</td>
<td>Low-Reactive Subjects</td>
</tr>
<tr>
<td>Total</td>
<td>2.35 (0.05)</td>
</tr>
<tr>
<td>Men</td>
<td>2.35 (0.07)</td>
</tr>
<tr>
<td>Women</td>
<td>2.36 (0.07)</td>
</tr>
</tbody>
</table>

Abbreviations: ES, effect size, Cohen d; OFC, orbitofrontal cortex; PFC, prefrontal cortex.
at 4 months of age. To our knowledge, there are no previous reports of a relation between infant temperament and brain structure in either infancy or adulthood. As summarized earlier, these temperamental differences have functional consequences lasting into adolescence.

LEFT OFC

We suggest that low-reactive subjects are able to modulate their hedonic tone in a more positive direction more effectively than high-reactive subjects because of more robust pathways in this subregion of the OFC that suppress unpleasant feelings. Functional neuroimaging studies support a central role for this subregion of the left OFC in hedonic processing54,55 and the reappraisal of negative emotion in a more positive direction.56 The posterior-lateral limb of the cluster may relate to a distinct pattern of heavy avoidance and defensive behaviors including a response we called arching of the back, a response seen almost exclusively in 4-month-old high-reactive infants. Direct projections to the hypothalamus from this subregion of the ventromedial prefrontal cortex can also activate the medulla and sympathetic chain, resulting in the increases in blood pressure and heart rate seen in inhibited children in response to the unfamiliar.

Furthermore, this region is reciprocally connected with the posterior parahippocampal gyrus72,73 and receives a unilateral projection from the hippocampus73,74 and hence may play an important role in detecting whether a person, place, or object is novel or familiar. A study of face perception showed greater activation of the right medial OFC, bilateral amygdalae, and right inferior parietal cortex when subjects viewed images of unfamiliar individuals, compared with viewing images of themselves.75 In that study, the maximum functional MRI activation in the right medial OFC occurred at precisely the same Talairach coordinates where we detected the largest temperament-related thickness difference. The flailing arms and legs characteristic of high-reactive infants in response to unexpected, or intense stimuli.78 Finally, the frequent distress vocalizations of high-reactive infants are mediated by direct projections from the medial prefrontal network to the periaqueductal gray and anterior cingulate.

These structural differences in the cerebral cortex of adults that correlate with infant temperament are present even when we excluded subjects with major depression or social phobia (Table 4). These findings therefore point to an early temperamental marker of vulnerability (or conversely resilience) to depressive and anxiety disorders. These anatomical features may represent novel endophenotypes for genetic analysis.

Several limitations merit comment. The thickness differences in the OFC and ventromedial cortex in these data are about 10% to 12%. Using similar techniques, regional thickness differences in the cerebral cortex of about 10% have been found in subjects with autism (including OFC)79 and 3% to 8% in patients with schizophrenia.80 The variations in thickness of the cortex we report could be poten-

### Table 4. Cortical Thickness by Temperament in Subjects Without Social Anxiety Disorder, MDD, and History of Medication Use

<table>
<thead>
<tr>
<th></th>
<th>Right Ventromedial PFC, mm, Mean (SEM)</th>
<th>Left OFC, mm, Mean (SEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High-Reactive Subjects</td>
<td>Low-Reactive Subjects</td>
</tr>
<tr>
<td>Total sample</td>
<td>2.35 (0.05)</td>
<td>2.09 (0.05)</td>
</tr>
<tr>
<td>Without social anxiety</td>
<td>2.33 (0.06)</td>
<td>2.10 (0.05)</td>
</tr>
<tr>
<td>Without MDD</td>
<td>2.39 (0.06)</td>
<td>2.09 (0.05)</td>
</tr>
<tr>
<td>Without medication use</td>
<td>2.38 (0.05)</td>
<td>2.09 (0.05)</td>
</tr>
</tbody>
</table>

Abbreviations: ES, effect size, Cohen d; MDD, major depressive disorder; OFC, orbitofrontal cortex; PFC, prefrontal cortex.
tially related to variation in the size or density of neurons, inhibitory interneurons, glial cells, or in the size and density of unmethylated neuronal processes (dendrites, dendritic spines, and axons) referred to as neuropil. The current state of high-resolution MRI cannot address which of these components contribute to the cortical thickness differences observed. Furthermore, because imaging data were not collected in infancy, these findings cannot address the question of whether the structural differences we report are primary and could be detected earlier or whether they develop over time because of genetic factors, environmental influences, or some interaction of the two. Understanding these developmental mechanisms could offer new avenues for the understanding of mood and anxiety disorders.

Submitted for Publication: May 20, 2008; final revision received March 25, 2009; accepted April 24, 2009.

Author Affiliations: Developmental Neuroimaging and Psychopathology Laboratory, Department of Psychiatry (Dr Schwartz, Stewart, and Snidman, Messrs Kunwar and Wallace, and Ms Moran, Viner, and Covino), and Department of Radiology (Dr Greve), Massachusetts General Hospital, Harvard Medical School (Drs Greve, Stewart, and Snidman), and Children’s Hospital (Dr Snidman), Boston, Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown (Drs Schwartz, Greve, and Vangel and Messrs Kunwar and Wallace), and Department of Psychology, Harvard University (Dr Kagan), General Clinical Research Centers, Massachusetts General Hospital (Dr Vangel), and Massachusetts Institute of Technology (Dr Vangel), Cambridge.

Correspondence: Carl E. Schwartz, MD, Massachusetts General Hospital, Psychiatric Neuroscience Program, Bldg 149, 13th Street, Charlestown, MA 02129 (carl_schwartz@hms.harvard.edu).

Financial Disclosure: None reported.

Support/ Funding: This study was supported by National Institutes of Mental Health grant 5R01MH071467 (Dr Schwartz), National Center for Research Resources grants 5P41 RR14075-07 (Center for Functional Neuroimaging Technologies), 5M01RR001066-27 (General Clinical Research Center), and 1 UL1 RR025758-01 (Harvard Clinical and Translational Science Center), the Mental Illness and Neuroscience Discovery Institute, and the Athinoula A. Martinos Center for Biomedical Imaging.

Previous Presentations: Portions of this work were presented at the Symposium on Biological Complexity: Genes, Circuits, and Behavior; sponsored by the Salk Institute, Nature, and Fondation IPSEN; January 10, 2008; La Jolla, California; and at the 47th Annual Meeting of the American College of Neuropsychopharmacology; December 7, 2008; Scottsdale, Arizona.

Additional Contributions: We thank the families and children who have stayed with the study for 18 years. With appreciation to Doreen Arcus, PhD, for participating in the classification of the 4-month-old infants, Dost Ongur, MD, PhD, for helpful discussions, and Grazyna Rajkowska, PhD, for confirming the correlation of our findings with sections from her postmortem studies of depression.