Supporting Information

Rauh et al. 10.1073/pnas.1203396109

SI Methods

MRI Acquisition. Anatomical images were acquired on a GE Signa 3 Tesla whole-body scanner with a body transmitter coil and an eight-channel head receiver coil. High-resolution, T1-weighted brain images were acquired using a fast spoiled gradient-recall sequence: inversion time = 500 ms, repetition time = 4.7 ms, echo time = 1.3 ms, field of view = 24 cm, image matrix = 256 × 256, acceleration factor = 2, number of slices = 160, slice thickness = 1 mm encoded for sagittal slice reconstruction, providing voxel dimensions of 0.9375 × 0.9375 × 1.0 mm.

Image Segmentation. All processing was performed on Sun Ultra 10 workstations using ANALYZE 8.0 Biomedical Imaging Resource (Mayo Foundation) and software developed in-house, while blind to subject characteristics and hemisphere (images were randomly flipped in the transverse plane before preprocessing). Morphometric analyses were performed with the MRI dataset resliced to correct for any residual head rotation, tilt, or flexion/extension.

Preprocessing. Large-scale variations in image intensity were corrected by an algorithm developed at the Montreal Neurological Institute (1). Extracerebral tissues were removed by an automated tool (2) that uses an anisotropic filter to smooth image intensity and a Marr–Hildreth edge detector (3) to identify 3D edges, before selecting as the brain the largest connected component with a closed boundary. Connecting dura was removed manually on each sagittal slice and checked in orthogonal views. The brainstem was transected at the pontomedullary junction.

Cortical gray matter segmentation. As briefly described in the text, gray-scale values of "pure" representations of cortical gray and white matter were sampled bilaterally at four standard locations throughout the brain (frontal, temporal, occipital, and parietal) using an 8 × 8 = 64 pixel array, sufficiently large to provide statistical stability but small enough to avoid partial volume effects that include other tissue types. These four values were then averaged for each tissue type, and a threshold value halfway between the mean gray and white matter values was computed for a slice in the imaging volume. We invoked these threshold values on a slice-by-slice basis to provide an initial rough classification of cortical gray and white matter throughout the cerebrum. This classification was hand edited in the coronal and transverse views to provide the most accurate segmentation possible of the cortical mantle. The intraclass correlation coefficient using a two-way random effects model (4) to measure reliability of cortical gray matter volumes was 0.98.

Overview of the Analysis of Surface Morphologies. The analysis of cerebral surfaces previously has been rigorously validated using synthetic and real-world datasets (5). Comparing surfaces either within one group or across two or more groups of brains requires determining which points on the surface of each brain correspond anatomically with the points on the surfaces of all of the other brains in the comparison. After establishing point correspondences, a signed distance (the Euclidean distance, with positive distances for outward deformations and negative distances for inward deformations) between the corresponding points within and between groups can be evaluated statistically. Our method for determining these point correspondences across individual brain surfaces uses a two-step procedure. In the first step, each brain is coregistered using a similarity (rigid body transformation with a global scaling) transformation to the template brain such that the cerebral surfaces are moved to a close approximation to the template surface. In the second step, each participant brain is treated as a fluid that is flowing into the template brain, and therefore constitutes a high-dimensional, nonlinear warping of the entire 3D volume of each participant brain to the entire 3D volume of the template brain. Each participant brain therefore appears exactly as the template brain appears, including every point of their outer surfaces, thereby establishing precise point-to-point correspondences across the surface of each participant brain with the template brain. The nonlinear warping of each participant brain is then reversed to bring each brain into the initial correspondence established by the similarity coregistration, bringing each label for correspondence matching of cerebral surfaces that had been established previously through the nonlinear warping. The physical distances are measured from each point on the surface of each participant's brain to the corresponding point on the surface of the template brain. Those distances at each point of the template brain, one for each participant, constitute a continuous variable that can be either compared across groups of participants, related to another variable within a group of participants, or simply quantified using traditional descriptive statistics. The parameters of those statistical analyses are then color-coded and displayed at each voxel on the surface of the template brain. Finally, measures of cortical thickness at each point on the surface of each participant's brain can replace distance from the template brain as the continuous variable of interest in those statistical analyses. The following sections provide details of the steps in coregistration, the procedures used in the selection of the template brain, and the statistical modeling procedures used in our population of participants.

Initial Similarity Coregistration. The random flips were first reversed to provide their original correct orientation. The brains were then brought into a common coordinate or template space. Following isolation of the brain from nonbrain tissue (above), we used mutual information of gray scale values to register the brain of each participant to the template brain (6). The template brain and each of the other brains in our sample were treated as rigid bodies as the optimal translation, rotation, and scaling parameters maximized the mutual information $I(F;R)$ between each brain and the template. Let $F = \{F_1, F_2, \ldots, F_n\}$ denote the random field from which one of the volumes, called the participant volume, was sampled. Similarly, let $R = \{R_1, R_2, \ldots, R_m\}$ be the random field from which the other volume, called the template volume, was sampled. Assuming that the random variables $\{F_1, F_2, \ldots, F_n\}$ were independently and identically distributed, let F be the random variable which represented the voxel intensities in the participant volume. Similarly, assuming that $\{R_1, R_2, \ldots, R_m\}$ were independently and identically distributed, let R be the random variable representing the voxel intensities in the template volume. Then the mutual information $I(F;R)$ between the two random variables F and R with marginal probability density functions $P_F(f)$ and $P_R(r)$ and joint probability density function $P_{F,R}(f,r)$ is defined to be $I(F,R) = \sum_{f,r} P_{F,R}(f,r)\, log_2\left(\frac{P_{F,R}(f,r)}{P_F(f)\times P_R(r)}\right)$ Let α denote the set of transformation parameters—three translations, three rotations, and scale—by which the participant volume was transformed. Then the probability density function of the transformed participant volume, as a function of α, is denoted as $P_{F,\alpha}(f)$. Similarly, the joint density function is denoted by$P_{F,R,\alpha}(f,r)$. The probability density function,$P_{R,\alpha}(r)$, is a function of α because we

were evaluating the mutual information of overlapping regions in the two volumes for various values of α. The joint and the marginal probability density functions were estimated using the Parzen window method (7). Using these notations, the mutual information, $I_\alpha(F,R)$, was evaluated as a function of α as:

$$I_\alpha(F,R) = \sum_{f,r} P_{F,R,\alpha}(f,r)\, log_2\left(\frac{P_{F,R,\alpha}(f,r)}{P_{F,\alpha}(f) \times P_{R,\alpha}(r)}\right). \quad \textbf{[S1]}$$

The optimal set of transformation parameters α that maximized $I_\alpha(F,R)$ were then found using the method of stochastic gradient descent (6).

Nonlinear Warping Using Fluid Flow Dynamics. Following the similarity transformation of each brain to the template brain described above, we used the method of fluid flow (FF) to refine this coregistration. The FF method uses a nonlinear deformation of one brain surface to another by treating the gray scale image of one brain as a fluid that flows into the gray scale image of another brain so as to match its gray scale characteristics precisely (8). The FF method therefore uses a high-dimensional, nonlinear warping of images to identify corresponding points on the participant and reference surfaces. This algorithm modeled the warping of gray scale in the images as the flow of fluid using the partial differential equations (PDEs) of the Navier–Stokes equations. In FF dynamics, nonlinear deformations of large magnitude are permitted because the stress (force per unit area) that restrains local deformations relaxes over time. The PDE governing the deformation is as follows (8):

$$\mu\nabla^2\vec{v}(\lambda+\mu)\vec{\nabla}(\vec{\nabla}\cdot\vec{v}) + \vec{b}(\vec{u}) = \vec{0}, \quad \textbf{[S2]}$$

where $\nabla^2 = \nabla^T\nabla$ is the Laplacian operator, $(\vec{\nabla}\cdot\vec{v})$ is the divergence operator, μ and λ are the viscosity coefficients, and $\vec{v}(\vec{x},t)$ is the velocity of the particle at time t and position $\vec{x}$ in the Eulerian reference frame. In the Eulerian reference frame, a voxel located at $\vec{x}$ at time t, with a displacement vector $\vec{u}(\vec{x},t)$, originated from a point $\vec{x} - \vec{u}(\vec{x},t)$. In Eq. **S2**, the term $\vec{b}(\vec{u(\vec{x},t)})$ denotes the body force acting on a particle located at $\vec{x}$ at time t, which is the driving force for the local nonlinear deformations of an image.

The PDE in Eq. **S2**, defined on a domain $\Omega = [0,1]^3$, is a boundary value problem that was solved numerically using the successive overrelaxation method with checker board updates (9). Zero boundary conditions were assumed for $\vec{v}(\vec{x},t)$ along the boundary $\partial\Omega$ of the domain Ω at all t, and therefore the displacements of the voxels at $\partial\Omega$ is zero. Because the displacement field $\vec{u}(\vec{x},t)$ was given in the Eulerian reference frame, the velocity field $\vec{v}(\vec{x},t)$ and the displacement fields $\vec{u}(\vec{x},t)$ were related as

$$\vec{v} = \frac{d\vec{u}}{dt} = \frac{\partial\vec{u}}{\partial t} + \sum_i v_i \frac{\partial\vec{u}}{\partial x_i}. \quad \textbf{[S3]}$$

We used the forward time centered space (9) method for the numerical estimation of the various partial derivatives in Eq. **S2**.

The modeling of fluid dynamics has been used extensively in various settings for the registration of 3D images (8). We used this algorithm, however, for identifying the corresponding points on the surface of the participant and template brains, rather than for simple registration. Using this high-dimensional warping algorithm, each participant brain was warped to the template brain, and points on the surface of the template brain were used to label points on the surface of each participant's brain. Each participant brain was then unwarped and brought back along with the labeled points to its original conformation in the common coordinate space established in the initial similarity transformation.

Measurement of Surface Distances from the Template Brain. Signed Euclidean distances between the corresponding labeled points of the template and participant brain were calculated. These distances were positive for outward deformation and negative for inward deformation of the surface of individual participants relative to the template surface. The distances computed at each point at the cortical surface were statistically analyzed using multiple linear regressions (below).

Measurement of Cortical Thickness. From the coregistered brain of each participant we subtracted its cortical mantle. We then used a 3D morphological operator to distance-transform this brain without the cortex from the coregistered brain of the same participant that contained the cortex (10, 11). This operation calculated cortical thickness as the smallest distance of each point on the external cortical surface from the outermost surface of the white matter in the coregistered brain. Because these thicknesses were measured in template space with the brain and its cortical thickness scaled during the similarity transformation, their values inherently accounted for generalized scaling effects within the cerebrum.

Selection of the Template Brain. The detection, localization, and interpretation of the statistically significant differences between groups of participants conceivably could depend on the choice of the template brain, because the estimated registration parameters and the established correspondences of points on the surfaces of the brain could depend upon the degree to which the selected template brain is representative of the population of participants being studied. Use of a synthetic, average brain can also be considered for the template. Creating a synthetic average, however, may not always be possible. For example, the pattern of gyri and sulci on the cortical surface vary significantly across individuals even within the healthy population (12). Thus, when not all participants have a particular gyrus or sulcus, how to generate a synthetic, average cortical surface is not clear. Moreover, averaging brain images across participants blurs boundaries between white and gray matter surfaces or between gray matter and CSF, especially at complex, convoluted surfaces such as those created by gyri at the surface of the brain, thereby increasing registration errors, increasing variance, and decreasing statistical power to detect real effects of interest at the surface of the brain. Therefore, we used a single representative brain as a template, rather than one derived by averaging brains across multiple people, because a single brain has well-defined tissue interfaces, including those at CSF/gray matter or gray/white matter interfaces.

We used a two-step procedure to select a brain as the template brain that was as representative as possible of the healthy participant sample being studied. First, one participant whose demographic characteristics was nearest the group average was selected preliminarily as the template brain. Second, the brains for all remaining healthy participants in the sample were registered to that preliminary template. The point correspondences across their surfaces were determined as described above and then the distances of those points from the corresponding points on the template surface were calculated. The brain for which all points across its surface were closest in the least-squares sense to the average of the distances across those points for the entire sample was selected to be the final template. The registration process, determination of point correspondences, and calculation of distances across surfaces were then repeated for all participants in the sample against this final template brain. The distances between the surfaces of the final template and the participant brains were then compared across individuals or groups.

Voxel-Wise Statistical Modeling Across the Cerebral Surface. To control for the effects of covariates (age, sex) on surface morphology, we performed a multiple-variable linear regression analysis (13) at each point on the reference surface:

$$d_i = \beta_0 + \beta_1 * Age + \beta_2 * Sex + \beta_3 * ADHD; \forall i = 1, \ldots, n,$$

where d_i was either (*i*) the set of signed Euclidean distances for the *i* th participant or (*ii*) the measure of cortical thickness for the *i* th participant at that voxel, and *n* is the number of participants in the entire sample. We computed the correlation β_3 between the distances and group membership (low or high exposure) and then computed the *P* value of this correlation using the Student *t* test. We did not covary for overall brain volume because each participant's brain was coregistered to the template brain.

Correction for Multiple Comparisons. We applied a method for false discovery rate (FDR) (14, 15) to limit type 1 errors when performing statistical analysis at each voxel across the surface of the brain. The multiple-variable linear regression was performed at each voxel on the surface of the brain, and if at a voxel the *P* value of the correlation was smaller than a specified significance level, then the null hypothesis is rejected. Because the null hypothesis was evaluated at thousands of voxels across the surface, a large number of null hypotheses would be rejected by chance. Although we can apply Gaussian random field-based methods to control for family-wise error rate (FWER), these methods are conservative and have low power to detect real effects because they control for any significant findings across the entire surface. A method for FDR, however, is a more powerful statistical method by allowing a prespecified number of false discoveries among all null hypotheses that are rejected. In our experiments, we set the FDR = 0.05. The *P* values of the voxels that survived the FDR correction were then color-encoded and plotted across the surface of the brain.

Sulcal Overlay. To aid the visual identification of the locations of significant findings on the surface of the brain, we overlaid onto the statistical maps of our template brain the sulcal boundaries previously identified on the International Consortium for Brain Mapping (ICBM) high-resolution, single-participant template (http://www.loni.ucla.edu/Atlases/Atlas_Detail.jsp?atlasid=5) (16). The sulcal boundaries and 3D labels of cortical gyri from the ICBM template were mapped onto the cortical surface of our template brain using our high-dimensional, nonrigid warping algorithm (above).

Controlling for Morphological Scaling Effects. Larger body size predicts larger overall brain size, and larger overall brain size predicts larger sizes of individual brain regions, a phenomenon termed "morphological scaling." The similarity transformation that we used for coregistration included a term for overall scale of the brain being registered to the template brain, which brings all brains in the sample to the same volume as the template brain and therefore controls for morphological scaling. To ensure that this control did not introduce unwanted changes in the structure of the data (i.e., in the average surface distances across exposure groups or in the variance of those measures at each voxel), we reversed the scaling transformation in the coregistered brains and plotted these same measures at each voxel, demonstrating similar structure in the scaled and unscaled data (Fig. S5). We also covaried for overall height of each participant in our regression models using the unscaled coregistered brains, but found that height overcorrected some of the surface distances in frontal and parietal regions, and it increased the variance in distance measures across most of the cerebral surface (Fig. S5, *Bottom*), indicating that covarying for height to control for scaling effects is undesirable in this dataset.

1. Sled JG, Zijdenbos AP, Evans AC (1998) A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 17:87–97.
2. Shattuck DW, Leahy RM (2002) BrainSuite: An automated cortical surface identification tool. *Med Image Anal* 6:129–142.
3. Marr D, Hildreth E (1980) Theory of edge detection. *Proc R Soc Lond B Biol Sci* 207: 187–217.
4. Shrout PE, Fleiss JL (1979) Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 86:420–428.
5. Bansal R, Staib LH, Whiteman R, Wang YM, Peterson BS (2005) ROC-based assessments of 3D cortical surface-matching algorithms. *Neuroimage* 24:150–162.
6. Wells WM, 3rd, Viola P, Atsumi H, Nakajima S, Kikinis R (1996) Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1:35–51.
7. Duda RO, Hart PE, Stork DG (2001) *Pattern Classification* (Wiley, New York), 2nd Ed.
8. Christensen GE, Joshi SC, Miller MI (1997) Volumetric transformation of brain anatomy. *IEEE Trans Med Imaging* 16:864–877.
9. Press WH (2002) *Numerical Recipes in C++: The Art of Scientific Computing* (Cambridge Univ Press, Cambridge, U.K.), 2nd Ed.
10. Haralick R, Shapiro L (1992) *Computer and Robot Vision* (Addison–Wesley, Vol 1, Chap 5).
11. Rosenfeld A, Pfaltz J (1968) Distance functions in digital pictures. *Pattern Recognit* 1 (1):33–61.
12. Ono M, Kubik S, Abernathey CD (1990) *Atlas of the Cerebral Sulci* (Thieme, New York).
13. Rosner B (2011) *Fundamentals of Biostatistics* (Brooks/Cole, Cengage Learning, Boston), 7th Ed.
14. Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188.
15. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300.
16. Mazziotta J, et al. (2001) A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos Trans R Soc Lond B Biol Sci* 356:1293–1322.

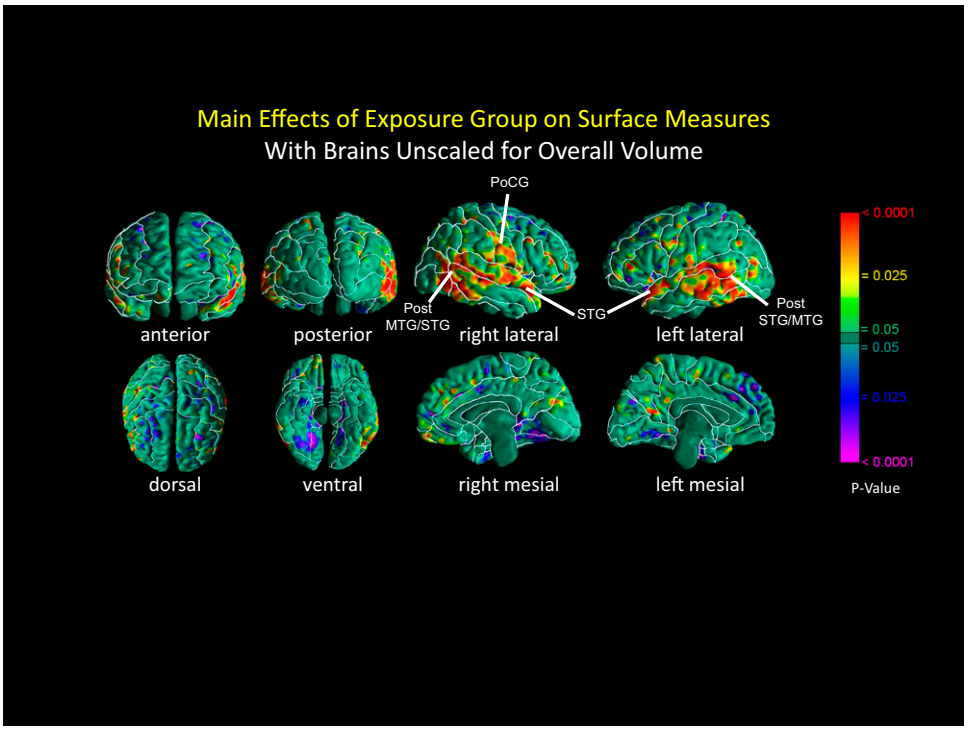

Fig. S1. Main effects of exposure group on surface measures with brains unscaled for overall volume. These analyses correspond to those shown in Fig. 1 of the main text, but without scaling for overall brain volume (i.e., without including scale in the similarity transformation used during procedures for image coregistration). Findings in the lateral temporal cortices bilaterally are essentially unchanged from findings in analyses that did include control for scaling effects (Fig. 1). MTG, middle temporal gyrus; PoCG, postcentral gyrus; STG, superior temporal gyrus.

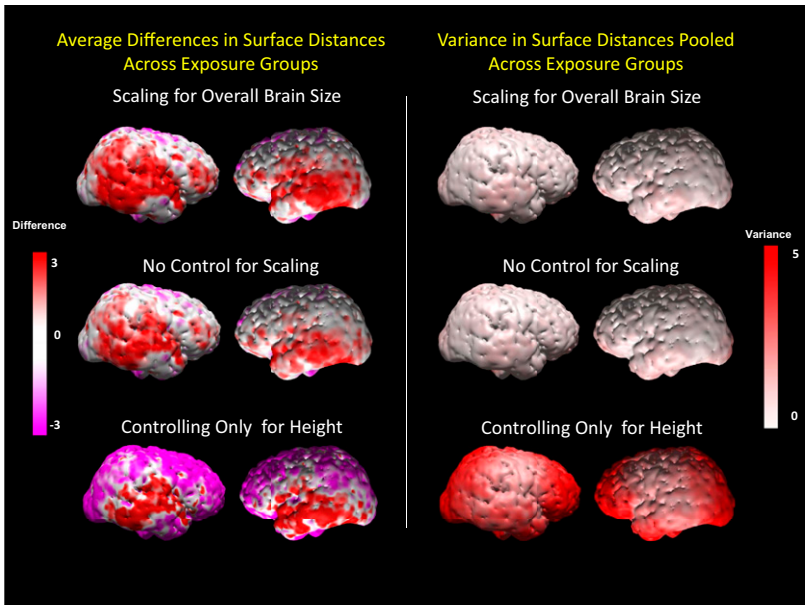

Fig. S2. Voxel-wise assessment of differing scaling effects on surface distances and variances. The influence of differing scaling effects on (*Left*) measures of surface distance (the distance from corresponding points on each participant's brain surface from the surface of the template brain) and (*Right*) the variance of that measure are plotted at each voxel for scaling that adjusts for overall brain size by (*i*) including scale as a parameter in the similarity transformation used to coregister each participant's brain with the template brain (*Top*), (*ii*) without including any adjustment for scaling effects in the brain (*Middle*), and (*iii*) when covarying measures of surface distance for each participant's height but without including brain size in the similarity transformation used to coregister each participant's brain to the template brain (*Bottom*). Surface distances and variances are highly similar with and without scaling for overall brain size. Scaling for height introduces some statistical artifacts into measures of surface distance along the dorsal frontal and parietal regions of the brain (purple) by overadjusting surface distances for scaling effects, and it increases the variance of surface distances across the entire cerebral surface, especially over the frontal, parietal, and temporal lobes, making height undesirable as a scaling covariate.

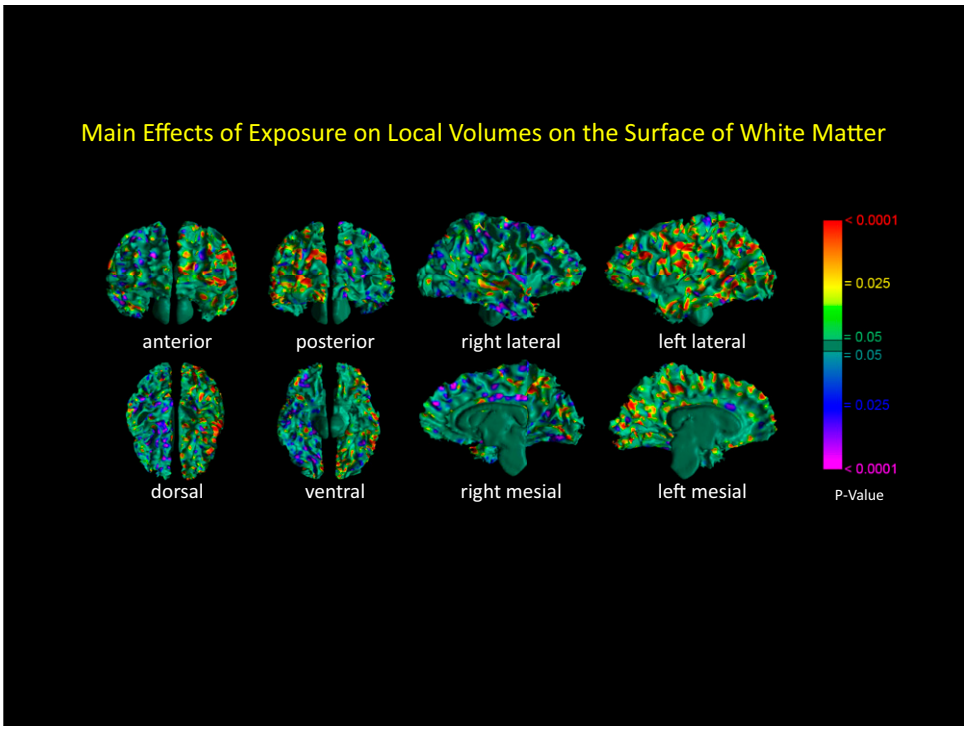

Fig. S3. Average surface measures of white matter in high- and low-chlorpyrifos (CPF) exposure groups. Shown here are color-coded maps comparing, across the high- and low-exposure groups, surface distances of white matter at each corresponding voxel of each participant's brain from the corresponding voxel of the white matter surface in the template brain. The pattern of differences across exposure groups is similar to the pattern of statistical significance of those differences depicted in the maps of *P* values comparing surface measures of the cerebral surface across groups (Fig. 1), particularly in the left hemisphere and in the vicinity of the Sylvian fissure of the left hemisphere. These analyses suggest that observed regional enlargement of the cerebral surface in the high-exposure group derived primarily from enlargement of underlying white matter.

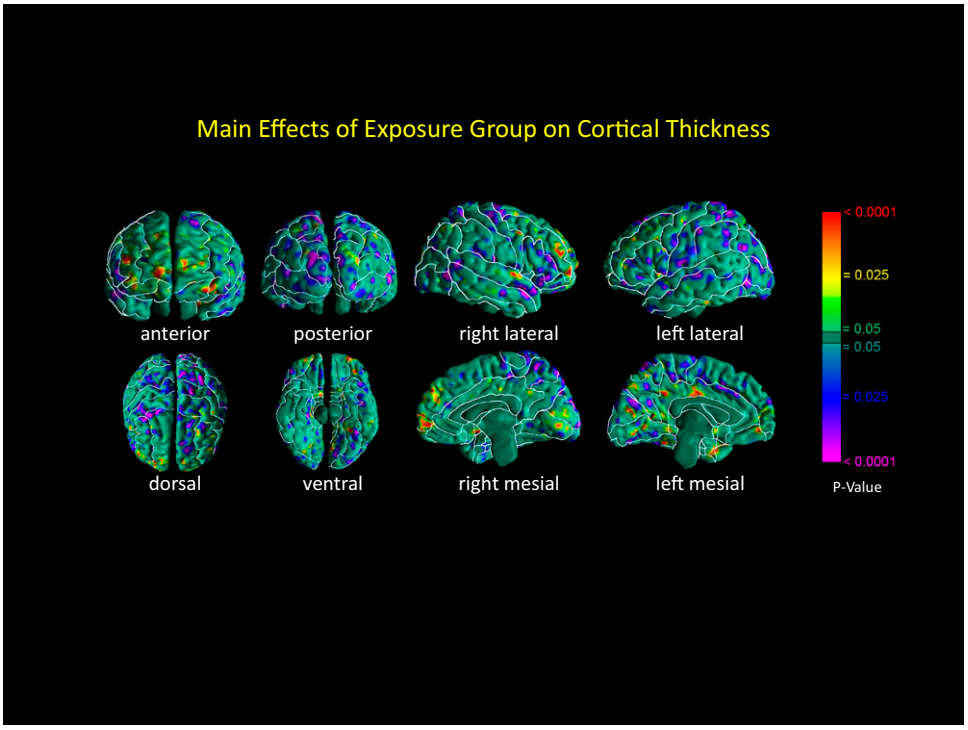

Fig. S4. Group differences in cortical thickness. Color-coded maps are shown for the statistical comparison of average differences in cortical thickness (in mm) between the high- and low-CPF exposure groups. Statistically significant group differences are small in spatial extent and scattered across the surface of the cortical mantle, but suggest the presence of reduced cortical thickness in the high-exposure group.

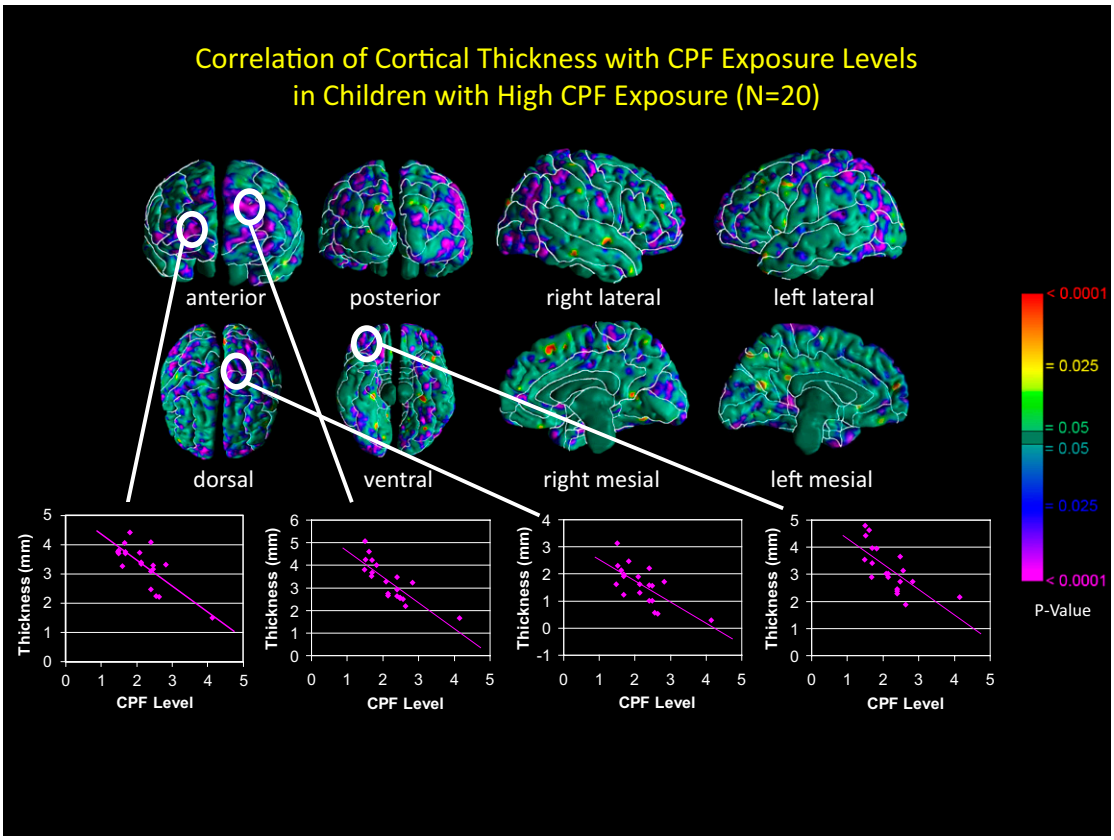

Fig. S5. Correlations of cortical thickness with CPF exposure levels. Within the high-CPF exposure group, cortical thickness correlated inversely with CPF exposure levels, such that higher exposure was associated with thinner cortices bilaterally in the dorsal parietal cortex, frontal pole, and orbitofrontal cortex. Scatterplots for the higher-exposure group show that cortical thickness declines with increasing level of CPF exposure in all regions within the high-exposure group.