SUPPLEMENTARY INFORMATION

Effects of Pollution Burden on Neural Function During Implicit Emotion Regulation and Longitudinal Changes in Depressive Symptoms in Adolescents

Uy *et al.*

Participants

Starting in 2013 and ending in 2017, 224 participants (N=132 females) were recruited from the San Francisco Bay Area for a longitudinal study of the effects of early life adversity on psychobiological functioning across adolescence. Inclusion criteria were that children were ages 9–13 years and were proficient in spoken English. Exclusion criteria were standard MRI contraindications (e.g., metal implants, braces), as well as a history of major neurological or medical illness or severe learning disabilities that would make it difficult to comprehend the study procedures and, for females, the onset of menses. Of the 224 participants, 187 completed an implicit emotion regulation task while undergoing an fMRI scan, 31 of whom were excluded from analyses (excessive motion: (N=28); not performing the task correctly (N=3)); 156 participants had at least one high-quality run/scan. Five of these participants did not provide addresses and, therefore, do not have environmental data, and six other participants did not report information on parental socioeconomic indicators. Of the remaining 145 participants, 142 reported depressive symptoms at baseline, 113 of whom also reported depressive symptoms four years later.

Exposures Indicators

The exposures component includes data "relating to pollution sources, releases, and environmental concentrations as indicators of potential human exposures to pollutants" (CalEnviroScreen 3.0 Report, p. 10). This component is comprised of 1) ozone concentrations in air; 2) PM 2.5 concentrations in air; 3) diesel particulate matter emissions; 4) drinking water contaminants; 5) use of certain high-hazard, high-volatility pesticides; 6) toxic releases from facilities; and 7) traffic density. Detailed descriptions of each indicator, adapted and quoted from the CalEnviroScreen 3.0 report, are provided below.

Ozone. Daily maximum 8-hour average concentrations for all monitoring sites in California were extracted from the California Air Resources Board (CARB)'s air monitoring network database for the summer months (May to October) for the years 2012-2014. Using Inverse Distance Weighting modeling, daily maximum 8-hour ozone concentrations for each tract were estimated from the center of each census tract and averaged across the summer months, which were averaged across the three years. Ozone values at census tracts with centers more than 50 km from the nearest monitor were not estimated using the model; for these tracts, the ozone value of the nearest air monitor was used (CalEnviroScreen 3.0 Report, pp. 22-23).

Particulate matter (PM) 2.5. "PM is a complex mixture of aerosolized solid and liquid particles including such substances as organic chemicals, dust, allergens, and metals. These particles can come from many sources, including cars and trucks, industrial processes, wood burning, or other activities involving combustion. The composition of PM depends on the local and regional sources, time of year, location, and weather" (CalEnviroScreen 3.0 Report, p. 26). PM 2.5 refers to particles that have a diameter of 2.5 micrometers or less. "PM2.5 annual mean monitoring data was extracted for all monitoring sites in California from CARB's air monitoring network database for the years 2012-2014. For all measurements in the time period, the mean concentrations were estimated at the geographic center of the census tract using a geostatistical method that incorporates the monitoring data from nearby monitors (ordinary kriging). Annual means were then computed for each year by averaging the quarterly estimates and then averaging those over the three-year period. PM2.5 values for census tracts with centers more than 50 km from the nearest monitor were assigned a concentration based on satellite observations for the years 2006-2012" (CalEnviroScreen 3.0 Report, pp. 27-28).

Diesel particulate matter. "Diesel PM is the particle phase of diesel exhaust emitted from diesel engines such as trucks, buses, cars, trains, and heavy-duty equipment. This phase is composed of a mixture of compounds, including sulfates, nitrates, metals, and carbon

particles" (CalEnviroScreen 3.0 Report, p. 32). The diesel PM indicator is distinct from the PM2.5 indicator and includes known carcinogens as well as particles in the ultrafine range (aerodynamic diameter less than 0.1 micrometer).

Diesel PM occurs from both on-road and off-road sources. "Gridded diesel PM emissions from on-road sources were calculated based on CARB's on-road emission model (EMFAC2013) to calculate 2012 county-wide estimates of diesel PM emissions for a July weekday in kg/day (http://www.arb.ca.gov/msei/modeling.htm). EMFAC2013 county-wide emission estimates are spatially distributed to 4-km-by-4-km grid cells based on the distribution of regional vehicle activity represented in local agency transportation networks and Caltrans' statewide transportation network (where local agency data are no available) using the Direct Travel Impact Model (DTIM4). Transportation networks are produced from travel demand modeling conducted by local agencies and Caltrans. Gridded diesel PM from non-road sources were calculated using county-wide estimates of diesel PM from non-road sources for a July weekend from CARB's emissions inventory forecasting system, CEPAM (http://www.arb.ca.gov/app/emsinv/fcemssumcat2009.php). County-wide emission estimates are spatially distributed to 4-km-by-4-km grid cells based on a variety of gridded spatial surrogate datasets. Each category of emissions is mapped to a spatial surrogate that generally represents the expected sub-county locations of source-specific activities. The surrogates include: lakes and coastline; population; housing and employment; industrial employment; irrigated cropland; unpaved roads; single-housing units; forest land; military bases; non-irrigated pasture land; rail lines; non-urban land; commercial airports; and ports. Diesel PM emissions estimates were adjusted to account for additional diesel PM emissions from sources on the Mexico side of the US-Mexico border. Resulting gridded emission estimates from the on-road and non-road categories were summed into a single gridded dataset. Gridded diesel PM emissions estimated were then allocated to census tracts in ArcMap using a weighted average

where the proportion of a grid-cell intersecting the populated portion (populated census blocks) of each census tract is used as the weight" (CalEnviroScreen 3.0 Report, pp. 33-34).

Drinking water contaminants. A drinking water contaminant metric was calculated for each census tract through 1) identification of drinking water system boundaries; 2) associating drinking water contaminant data with each water system and calculating average concentrations for each contaminant and system, and 3) re-allocating the average water contaminant concentration from the system boundaries to census tracts. Monitoring data for a subset of contaminants tested in drinking water across California were obtained from CDPH's Water Quality Monitoring database from 2005-2013. These contaminants included: arsenic, cadmium, hexavalent chromium, dibromochoropropane, lead, nitrate (NO3), perchlorate, radium 226 and radium 228, total trihalomethanes, tetrachloroethylene, trichloroethylene, 1,2,3-trichloropopanne, and uranium. "Time-weighted average concentrations of each contaminant were calculated for each year for each sample source within a system. The average yearly concentrations were then averaged to create a source concentration. Then, the source concentrations within a system were averaged to calculate one concentration value for each chemical in each system. Areas without system or sample source data were assigned the average groundwater quality data for sources in the township in which they were located (raw or untreated community or non-community water system data, Domestic Well Project water quality data, and Priority Basin water quality data). People in these areas were assumed to drink groundwater" (CalEnviroScreen 3.0 Report, p. 40).

Pesticide use. "Production agricultural pesticide use records were obtained for the entire state for the years 2012, 2013, and 2014. Production pesticide use (total pounds of selective active ingredient, filtered for hazard and volatility) for Meridien-Township-Range-Selection (MTRS) records were matched to census tracts and divided by each census tract's area" (CalEnviroScreen 3.0 Report, pp. 47-48).

Toxic releases from facilities. "The US Environmental Protection Agency (EPA) maintains a toxic substance inventory of on-site releases to air, water, and land, and underground injection of any classified chemical, as well as quantities transferred off-site, that are reported by each facility. US EPA has a computer-based screening tool called Risk Screening Environmental Indicators (RSEI) that analyzes these releases and models potential toxic exposures" (CalEnviroScreen 3.0 Report, p. 53). The Toxic Release Inventory (TRI) maintains a database of emissions and other release for certain toxic chemicals that is updated annually. Census tract-level estimates of "toxicity-weighted concentrations of modeled chemical releases to air from facility emissions and off-site incineration (averaged over 2011 to 2013) were made by taking a land-area weighted average of the block-level values for each tract. Census tracts were ordered based on their scores and were assigned percentiles" (CalEnviroScreen 3.0 Report, p. 56).

Traffic density. "Traffic density is used to represent the number of mobile sources in a specified area, resulting in human exposures to chemicals that are released into the air by vehicle exhaust, as well as other effects related to large concentrations of motor vehicles" CalEnviroScreen 3.0 Report, p. 59). A 150-meter buffer was placed around each census tract to account for roadways near tract boundaries. "The traffic volume data (from TrafficMetrix) was linked to the corresponding road segment (from TeleAtlas) in a geographic information system (GIS). The buffered census tracts were intersected with the linked data on traffic volumes and roads. For each road within the buffer, a length-adjusted volume was calculated and summed for all roads in the buffer. The total amount of road length within the buffered census tract was also calculated. Due to differences in the length of road segments across the state, a length-adjusted traffic volume metric was selected. This metric multiplies traffic volumes by the length of the road segment. Traffic density was then calculated by dividing the sum of all length-adjusted traffic volumes within the buffered census tract (vehicle-km/hr) by the sum of the length of all road segments within the buffered census tract (km). Traffic density (vehicles-km/hr/km) is

represented as the number of vehicles (adjusted by road segment lengths in kilometers) per hour per kilometer of roadways within the buffered census tract for 2013" (CalEnviroScreen 3.0 Report, p. 61).

Environmental Effects Indicators

The environmental effects component includes "adverse environmental conditions caused by pollutants" (CalEnviroScreen 3.0 Report, p. 10). This component represents "the presence of pollutants in a community, rather than exposure to pollutants (CalEnviroScreen 3.0 Report, p. 13). Environmental effects include environmental degradation, ecological effects, and threats to the environment and communities, which can have harmful effects on different components of the ecosystem (e.g., limiting people's use of ecosystem resources such as fishing or swimming). Indicators in the environmental effects component include 1) toxic cleanup sites; 2) groundwater threats from leaking underground storage sites and cleanups; 3) hazardous waste facilities and generators; 4) impaired water bodies; and 5) solid waste sites and facilities. Detailed descriptions of each of these indicators, adapted from the CalEnviroScreen 3.0 report, are presented below.

Toxic cleanup sites. "Data on cleanup site type, status, and location for the entire state were obtained from Department of Toxic Substances Control (DTSC)'s EnviroStor database (downloaded on December 2016) and mapped and geocoded in ArcMap. Polygon boundaries of California National Priority List sites were identified and sites were assigned a score of 10 or 12, indicating a federal Superfund site. Remaining sites were scored on a weighted scale of 0 to 12 in consideration of both the site type and status. Higher weights were applied to Superfund, State Response sites, cleanups, and sites undergoing active remediation and oversight by DTSC, relative to those with little or no state involvement. The weights for all sites were adjusted based on the distance they fell from populated census blocks. Sites further than 1000m from any populated census block were excluded from the analysis. Site weights were

adjusted by multiplying the weight by 1 for sites less than 250m, 0.5 for sites 250-500m, 0.25 for sites 500-750m, and 0.1 for sites 750-1000m from the nearest populated census blocks within a given tract. Each census tract was scored based on the sum of the adjusted weights (in ArcMap)" (CalEnviroScreen 3.0 Report, pp. 67-68).

Groundwater threats. The storage and disposal of hazardous materials on land and in underground storage tanks at various types of commercial, industrial, and military sites pose threats to groundwater quality through the leakage of hazardous substances that degrade soil and groundwater and increase volatile contaminants in the air. Data on cleanup site type, status, and location for the entire state were obtained from GeoTracker (downloaded on December 2016) and mapped and geocoded. Sites that were undergoing permit process (proposed) and those no longer accepting waste (closing/closed), which were not indicative of a hazard or environmental risk, were excluded. "Each remaining site was scored on a weighted scale of 1 to 15 in consideration of both the site type and status, which were then adjusted based on the distance they fell from populated census blocks. Site weights were adjusted by multiplying the weight by 1 for sites less than 250m, 0.5 for sites 250-500m, 0.25 for sites 500-750m, and 0.1 for sites 750-1000m from the nearest populated census blocks within a given tract. Sites outside of a census tract, but less than 1000m from one of that tract's populated blocks were similarly adjusted based on the distance to the nearest block from that tract. Each census tract was scored based on the sum of the adjusted weights for sites it contains or is near" (CalEnviroScreen 3.0 Report, pp. 74-75).

Hazardous waste generators and facilities. "Most hazardous waste must be transported from hazardous waste generators to permitted recycling, treatment, storage, or disposal facilities by registered hazardous waste transporters" (CalEnviroScreen 3.0 Report, p. 82). These processing and disposal of hazardous waste may contaminate nearby air, water, and soil. Permitted facility data were obtained from the Department of Toxic Substances Control (DTSC) website (downloaded December 2016). "Facilities were scored on a weighted scale in

consideration of the type and permit status for the facility and then mapped or geocoded. Hazardous waste generator data were obtained from DTSC from the Hazardous Waste Tracking System for 2012 to 2014 and only large quantity generators (producing over 1,000 kg of waste per month2 for at least one of the three years) and generators producing RCRA waste were included. Facilities were scored on a weighted scale in consideration of the volume of waste generated. Site locations were then mapped or geocoded and the weights for facilities were adjusted based on the distance they fell from populated census blocks. Site weights were adjusted by multiplying the weight by 1 for facilities less than 250m, 0.5 for sites 250-500m, 0.25 for sites 500-750m, and 0.1 for sites 750-1000m from the nearest populated census blocks within a given tract" (CalEnviroScreen 3.0 Report, pp. 83-84).

Impaired water bodies. "Contamination of California streams, rivers, and lakes by pollutants can compromise the use of the water body for drinking, swimming, fishing, aquatic life protection, and other beneficial uses, making such bodies 'impaired.' Data on water body type, water body identification, and pollutant type for 2012 were downloaded from the State Water Resources Control Board website and all water bodies were identified in all census tracts. The number of pollutants listed in streams or rivers that fell within 1 kilometer (km) or 2 km of a census tract's populated blocks were counted. The 2 km buffer distance was applied to major rivers (>100 km in length, plus the Los Angeles River and Imperial Valley canals and drainage ways). The 1 km buffer distance was applied to all smaller streams/rivers. The number of pollutants listed in lakes, bays, estuaries or shoreline that fell within 1 km or 2 km of a census tract's populated blocks were counted. The 2 km buffer distance was applied to major lakes or bays greater than 25 square kilometers in size, plus all the Sacramento/San Joaquin River Delta waterways. The 1 km buffer distance was applied for all other lakes/bays. These pollutant counts were summed for every census tract and each census tract was scored based on the sum of the number of individual pollutants found within and/or bordering it. For example, if two

stream sections within a census tract were both listed for the same pollutant, the pollutant was only counted once" (CalEnviroScreen 3.0 Report, pp. 89-91).

Solid waste sites and facilities. The California Department of Resources Recycling and Recovery (CalRecycle) maintains data on waste facilities that operate within the state, including abandoned, closed, or illegal sites. Closed, illegal, and abandoned waste sites as of December 2016 were scored on a weighted scale based on CalRecycle's prioritization categories. Active solid waste sites as of December 2016 were scored on a weighted scale based on the type of solid waste operation (e.g., solid waste landfill, construction, composting). "The weights for all sites, including the large landfill perimeters, were adjusted based on the distance they fell from populated census blocks. Site weights were adjusted by multiplying the weight by 1 for sites less than 250m, 0.5 for sites 250-500m, 0.25 for sites 500-750m, and 0.1 for sites 750-1000m from the nearest populated census blocks within a given tract. Odor complaints regarding composting facilities are commonly made more than 1000 m from these facilities. Because of this concern the buffer distances (and site weights) for composting sites were adjusted as follows: 1 for sites less than 500m, 0.5 for sites 500 – 1000m, 0.25 for sites 1000 – 1500m, and 0.1 for sites 1500 – 2000m from the nearest populated census blocks within a given tract. Each census tract was scored based on the sum of the adjusted weights for sites it contains or is near" (CalEnviroScreen 3.0 Report, p. 96).

Pollution Burden Percentile

For each indicator, the indicator values for the census tracts for the entire state were ordered from highest to lowest and a percentile value was assigned to each census tract. The percentiles for all the individual indicators in each component were then averaged, representing the score for that component. When combining the exposures and environmental effects components to calculate pollution burden scores, the environmental effects score was half-weighted because CalEnviroScreen determined the "contribution to possible pollutant burden

from the environmental effects component to be less than those from sources in the exposures component" (CalEnviroScreen 3.0 Report, page 13-14). Pollution burden scores for each census tract were divided by the maximum value observed in the state and then multiplied by 10. These scores were then ordered from highest to lowest and a percentile score was calculated.

Socioeconomic Indicators

Educational Attainment. Data from the 2011-2015 American Community Survey (ACS) estimates was used to calculate the proportion of the population with less than a high school education for each census tract.

Housing Burdened Low-Income Households. Data from the 2009-2013 Housing and Urban Development Comprehensive Housing Affordability Strategy were used to estimate the proportion of households per tract with household incomes less than 80% of the county median and renter or homeowner costs that exceed 50% of household income.

Poverty. Data from the 2011-2015 ACS was used to calculate the proportion of individuals in each tract living below 200 percent of the federal poverty line.

Unemployment. Data from the 2011-2015 ACS was used to calculate, per tract, the "percent of the population over the age of 16 that is unemployed and eligible for the labor force. [This measure] excludes "retirees, students, homemakers, institutionalized persons except prisoners, those not looking for work, and military personnel on active duty" (CalEnviroScreen, p. 143).

MRI Acquisition

Neuroimaging data were acquired at the Stanford Center for Cognitive and Neurobiological Imaging using a 3T General Electric (GE) Discovery MR750 scanner with a 32-channel head coil (Nova Medical). We acquired a structural T1-weighted IR-prep, fast SPGR

(3D BRAVO) sequence (repetition time [TR]=6.3ms; echo time [TE]=3.1ms; flip angle=12°; matrix=256 x 256; FOV=23.04cm, slice thickness=0.9mm, voxel resolution=0.9mm isotropic, acquisition time=5:16) and two runs of the functional MRI task using T2*-weighted gradient echo EPI pulse sequences (TR=2000ms; TE=30ms; flip angle=77°; matrix=70 x 70; FOV=22.4cm, slice thickness=3mm, voxel resolution=3.2mm x 3.2mm x 3.0mm, acquisition time=5:54).

MRI Preprocessing

Anatomical T1w images were first processed with FreeSurfer (v.6.0.1) (1) for surface-based registration, followed by functional processing using fMRIPrep (v.20.2.1) (2,3). A detailed description of the fMRIPrep protocol is described below in the boilerplate generated from fMRIPrep. In short, the pipeline generated 1) two runs of fMRI data corrected for susceptibility-distortion artifacts and spatially registered to MNI standard space using a pediatric-specific template for individuals ages 7.5-13.5 years (4), which were then transformed to adult MNI152 space; and 2) confound regressors used in individual-level modeling.

fMRIPrep Methods Boilerplate

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 20.2.1 (Esteban, Markiewicz, et al. (2018); Esteban, Blair, et al. (2018); RRID:SCR_016216), which is based on *Nipype* 1.5.1 (Gorgolewski et al. (2011); Gorgolewski et al. (2018); RRID:SCR_002502).

Anatomical data preprocessing

A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset. The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al. 2010), distributed with ANTs 2.3.3 (Avants et al. 2008, RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the antsBrainExtraction.sh workflow (from ANTs), using MNIPediatricAsym:cohort-4 as target

template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823, Zhang, Brady, and Smith 2001). Volume-based spatial normalization to two standard spaces (MNIPediatricAsym:cohort-4, MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization: *MNI's unbiased standard MRI template for pediatric data from the 4.5 to 18.5y age range* [(**???**), RRID:SCR_008796; TemplateFlow ID: MNIPediatricAsym:cohort-4], *ICBM 152 Nonlinear Asymmetrical template version 2009c* [Fonov et al. (2009), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym],

Functional data preprocessing

For each of the 8 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. A B0-nonuniformity map (or *fieldmap*) was directly measured with an MRI scheme designed with that purpose (typically, a spiral pulse sequence). The *fieldmap* was then co-registered to the target EPI (echo-planar imaging) reference run and converted to a displacements field map (amenable to registration tools such as ANTs) with FSL's fugue and other *SDCflows* tools. Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using flirt (FSL 5.0.9, Jenkinson and Smith 2001) with the boundary-based registration (Greve and Fischl 2009) cost-function. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, Jenkinson et al. 2002). BOLD runs were slice-

time corrected using 3dTshift from AFNI 20160207 (Cox and Hyde 1997, RRID:SCR_005927). The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in MNIPediatricAsym:cohort-4 space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (2002)). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al. 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, Behzadi et al. 2007). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by thresholding the corresponding partial volume map at 0.05, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation).

Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al. 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer). First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Susceptibility distortion correction (SDC) was omitted. The BOLD reference was then co-registered to the T1w reference using flirt (FSL 5.0.9, Jenkinson and Smith 2001) with the boundary-based registration (Greve and Fischl 2009) cost-function. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, Jenkinson et al. 2002). BOLD runs were slice-time corrected using 3dTshift from AFNI 20160207 (Cox and Hyde 1997, RRID:SCR_005927). The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying the transforms to correct for head-motion. These resampled BOLD time-

series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in MNIPediatricAsym:cohort-4 space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (2002)). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al. 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, Behzadi et al. 2007). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by thresholding the corresponding partial volume map at 0.05, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped

from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al. 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.6.2 (Abraham et al. 2014, RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

Copyright Waiver

References

- Abraham, Alexandre, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gael Varoquaux. 2014. "Machine Learning for Neuroimaging with Scikit-Learn." *Frontiers in Neuroinformatics* 8. https://doi.org/10.3389/fninf.2014.00014.

- Avants, B.B., C.L. Epstein, M. Grossman, and J.C. Gee. 2008. "Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and

Neurodegenerative Brain." *Medical Image Analysis* 12 (1): 26–41. https://doi.org/10.1016/j.media.2007.06.004.

- Behzadi, Yashar, Khaled Restom, Joy Liau, and Thomas T. Liu. 2007. "A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI." *NeuroImage* 37 (1): 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042.

- Cox, Robert W., and James S. Hyde. 1997. "Software Tools for Analysis and Visualization of fMRI Data." *NMR in Biomedicine* 10 (4-5): 171–78. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

- Esteban, Oscar, Ross Blair, Christopher J. Markiewicz, Shoshana L. Berleant, Craig Moodie, Feilong Ma, Ayse Ilkay Isik, et al. 2018. "FMRIPrep." *Software*. Zenodo. https://doi.org/10.5281/zenodo.852659.

- Esteban, Oscar, Christopher Markiewicz, Ross W Blair, Craig Moodie, Ayse Ilkay Isik, Asier Erramuzpe Aliaga, James Kent, et al. 2018. "fMRIPrep: A Robust Preprocessing Pipeline for Functional MRI." *Nature Methods*. https://doi.org/10.1038/s41592-018-0235-4.

- Fonov, VS, AC Evans, RC McKinstry, CR Almli, and DL Collins. 2009. "Unbiased Nonlinear Average Age-Appropriate Brain Templates from Birth to Adulthood." *NeuroImage* 47, Supplement 1: S102. https://doi.org/10.1016/S1053-8119(09)70884-5.

- Gorgolewski, K., C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. Ghosh. 2011. "Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python." *Frontiers in Neuroinformatics* 5: 13. https://doi.org/10.3389/fninf.2011.00013.

- Gorgolewski, Krzysztof J., Oscar Esteban, Christopher J. Markiewicz, Erik Ziegler, David Gage Ellis, Michael Philipp Notter, Dorota Jarecka, et al. 2018. "Nipype." *Software*. Zenodo. https://doi.org/10.5281/zenodo.596855.

- Greve, Douglas N, and Bruce Fischl. 2009. "Accurate and Robust Brain Image Alignment Using Boundary-Based Registration." *NeuroImage* 48 (1): 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060.

- Jenkinson, Mark, Peter Bannister, Michael Brady, and Stephen Smith. 2002. "Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images." *NeuroImage* 17 (2): 825–41. https://doi.org/10.1006/nimg.2002.1132.

- Jenkinson, Mark, and Stephen Smith. 2001. "A Global Optimisation Method for Robust Affine Registration of Brain Images." *Medical Image Analysis* 5 (2): 143–56. https://doi.org/10.1016/S1361-8415(01)00036-6.

- Lanczos, C. 1964. "Evaluation of Noisy Data." *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis* 1 (1): 76–85. https://doi.org/10.1137/0701007.

- Power, Jonathan D., Anish Mitra, Timothy O. Laumann, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen. 2014. "Methods to Detect, Characterize, and Remove Motion Artifact in Resting State fMRI." *NeuroImage* 84 (Supplement C): 320–41. https://doi.org/10.1016/j.neuroimage.2013.08.048.

- Satterthwaite, Theodore D., Mark A. Elliott, Raphael T. Gerraty, Kosha Ruparel, James Loughead, Monica E. Calkins, Simon B. Eickhoff, et al. 2013. "An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data." *NeuroImage* 64 (1): 240–56. https://doi.org/10.1016/j.neuroimage.2012.08.052.

- Tustison, N. J., B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. 2010. "N4ITK: Improved N3 Bias Correction." *IEEE Transactions on Medical Imaging* 29 (6): 1310–20. https://doi.org/10.1109/TMI.2010.2046908.

- Zhang, Y., M. Brady, and S. Smith. 2001. "Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm." *IEEE Transactions on Medical Imaging* 20 (1): 45–57. https://doi.org/10.1109/42.906424.

Individual-Level Modeling of fMRI Data

fMRIPrep outputs were analyzed using FSL's (v. 6.0.1) *FEAT* (5,6). The first five volumes of each dataset were removed to allow for the stabilization of the longitudinal magnetization. The data were spatially smoothed (FWHM=5mm) and underwent high-pass temporal filtering (100.0s). Individual-level modeling was conducted using FSL's FMRIB's Improved Linear Model with pre-whitening to improve estimation of each voxel's time series. The voxel-wise general linear model (GLM) included regressors for each of the five task conditions and selected confounds generated from fMRIPrep: framewise displacement (FD), head-motion (six rigid-body parameters, temporal derivatives, and quadratic terms), and six anatomical components derived from component-based noise correction (CompCor) (7). Motion outliers (frames > 0.5mm FD or 1.5 DVARS) were censored by adding single-TR regressors at censored volumes. We excluded scans with more than 20% of volumes marked as outliers (n=28; mean FD = 0.16mm, SD=0.10). A double-gamma convolution was used to model the hemodynamic response function. Runs were combined using a fixed-effects model using *FLAME1* (FMIRB's Local Analysis of Mixed Effects; (8).

Group-Level Whole-Brain Analyses

To examine whether pollution burden was associated with neural activation during implicit emotion regulation and whether effects differed by affective valence, we conducted group-level whole-brain GLM analyses using FLAME1 on the condition x valence contrast (i.e., [*Negative Label > Negative Match*] > [*Positive Label > Positive Match*]) with pollution burden percentile scores (mean-centered) entered as a regressor in the GLM. Participant sex (-1=male, 1=female), age (mean-centered), household socioeconomic disadvantage (mean-centered),

neighborhood socioeconomic disadvantage (mean-centered), time difference between pollution measurement and scan date (mean-centered), and mean FD (mean-centered) were included as covariates. Z (Gaussianized T) statistic images were thresholded at Z>3.1, with a corrected cluster significance threshold of p<.025 (Bonferroni-corrected alpha=0.05/2 tests [whole-brain activation and PPI]). Anatomical localization of each cluster was defined using the FSL Harvard-Oxford probabilistic atlas. Following up a significant interaction of condition, valence, and pollution burden, we extracted parameter estimates of brain activation in the significant clusters from the *Label > Match* contrasts for each valence and conducted simple effects analyses for each valence to determine how pollution burden was related to implicit emotion regulation (i.e., *Label > Match*) of positively- vs. negatively-valenced stimuli.

Main Effects of Task

Whole-brain analyses examining the main effect of task conditions indicated that matching (relative to labeling) emotional stimuli activated visual cortex regions, amygdala, precentral, superior, and middle frontal regions that did not differ between positively- and negatively-valenced stimuli (Tables S4; Figure S3). Labeling (relative to matching) emotional stimuli activated temporal cortex regions, superior, inferior, and orbito- frontal regions, and the cingulate, and cerebellum, with greater activation in paracingulate and dorsolateral PFC in the regulation of positively- than of negatively-valenced stimuli (Table S5; Figure S4).

Whole-Brain GLM Results

Whole-brain analyses indicated that pollution burden was associated with bilateral medial prefrontal cortex (MPFC) activation for the condition x valence contrast (i.e., [*Negative Label > Negative Match*] > [*Positive Label > Positive Match*]) (left MPFC: -14, 56, 10, Z max = 4.11, 126 voxels; right MPFC: 14, 36, 6, Z max = 3.77, 146 voxels; Figure 1A). To probe the nature of the interaction, parameter estimates of left and right MPFC clusters were extracted from the *Label > Match* contrasts for each valence. The associations of pollution burden with left and right MPFC do not differ in slope (b=0.07, p=.474; Table S3 & Figure S5); consequently, we

averaged the MPFC parameter estimates across hemispheres in subsequent analyses. Further, visual inspection of the scatterplots depicting the association between pollution burden with bilateral MPFC by valence suggests that pollution burden was positively associated with MPFC activation for the *Negative Label > Negative Match* contrast, but not for the *Positive Label > Positive Match* contrast (Figure 1B and 1C). That is, greater pollution burden was associated with greater bilateral MPFC activation during implicit affective regulation of negatively-, but not of positively-, valenced stimuli.

Exploratory Analyses of Environmental Indicators

We conducted exploratory analyses to determine which specific environmental indicator(s) moderated associations between MPFC-DMN connectivity during implicit regulation of negatively-valenced stimuli and slopes of depressive symptoms. We conducted least absolute shrinkage and selection operator (LASSO) regression analyses, entering all of the environmental indicators (z-scored), their interactions with MPFC-DMN connectivity (z-scored), and all covariates (z-scored) simultaneously as predictors to determine whether there was a combination of environmental indicators that moderated the association between MPFC-DMN connectivity and the slope of depressive symptoms. LASSO is a regression analysis that performs both feature selection and regularization to select a subset of variables that are most strongly predictive of outcomes while shrinking non-predictive variables to zero to prevent overfitting. This approach has recently been used to identify factors that predict borderline personality disorder, depression, and conduct disorders in adolescents (9). We used the *glmnet* package in R to perform the LASSO. Similar to Beeney et al.'s (2021) procedure, we randomly split data into a training set (70%) and testing set (30%). We selected the lambda value that minimizes the mean square error (MSE) of the training set via cross-validation using 10-fold cross-validation and then used this lambda (0.0265) to run LASSO on the test set. The model retained seven interaction terms that best predicted slopes of depressive symptoms, including

most covariates (Table S5). Based on results from the LASSO regression, the environmental indicators that moderated the association between MPFC-DMN connectivity during implicit regulation of negative emotions and slopes in depressive symptoms were ambient ozone, diesel PM, toxic release from facilities, and living near cleanup sites, groundwater threats, hazardous waste facilities, and impaired water bodies.

Sensitivity Analyses: Testing for Residual Confounding

Although we controlled for variables that could be confounds in our models, there still may be effects of residual confounding. While there is little we can do to reduce the effects of unmeasured confound variables (10) or measurement error (11), one source of residual confounding can stem from mis-specifying the confound variables in our models (e.g., assuming that there are linear associations between confound variables and the outcome variable when the associations may, in fact, be non-linear) (12). Thus, we examined the bivariate plots of each continuous confound variable (i.e., age at scan, household socioeconomic disadvantage, neighborhood socioeconomic disadvantage, time difference between pollution measurement and scan dates, and mean FD) with each outcome variable (i.e., MPFC activation, MPFC-DMN connectivity, MPFC-IFG connectivity, depressive symptoms intercept, and depressive symptoms slope) to check for non-linearity (Figure S6A-E). Examination of the plots indicated that, except for a non-linear association between depressive symptoms intercept and slope (Figure S6E), modeling covariates using linear terms in the models is appropriate. Thus, we also conducted brain-behavior analyses predicting depressive symptoms slopes including the quadratic term of depressive symptoms intercept in the models. Analyses indicated that including the quadratic term of initial levels of depressive symptoms did not change our findings (Table S8).

Sensitivity Analyses: Clustering Participants by Census Tract

There were 125 unique census tracts in our sample of 145 people, indicating that 20 people resided in the same census tract with at least one other participant in the study. Therefore, we reran our analyses in a multilevel modeling framework with census tract as the

unit of analysis. Results indicated that using census tract as a grouping variable did not change the results (Table S9). Furthermore, the ICC was 0.04, suggesting that the correlations among observations within census tracts are low and that it is not necessary to cluster participants by census tract in our analyses.

Sensitivity Analyses: Including Early Life Adversity as a Covariate

Early Life Adversity. We assessed participants' exposure to early life adversity (ELA) through a structured interview using a modified version of the Traumatic Events Screening Inventory for Children (TESI-C; Ford et al., 2002). Using the TESI, we asked participants if they have experienced more than 30 types of life events, including direct exposure to or witnessing of severe accidents, illness or disaster, family or community conflict or violence, and sexual molestation. Each type of event endorsed by a participant was followed up with questions to obtain a deeper characterization of the experience. For example, participants were asked if they have "ever been in a really bad accident, like a car accident, a fall, or fire, where you or someone else could have been (or actually was) badly injured or killed." If this item was endorsed, participants were then asked how many times such an event happened (if multiple times, each event was recorded as a separate event), when the event happened, and whether and how badly they or someone else were really hurt. After the interview, the administrator then presented each event to a panel of three trained coders who, blind to the participants' subjective severity ratings, used a modified version of the UCLA Life Stress Coding System (13,14) to rate the objective severity of each event on a scale ranging from 0 (not impactful) to 4 (extremely severe and impactful), with half-point increments (inter-rater ICC=0.99). If there were any discrepancies among coders in ratings for an event, the coders discussed their ratings to arrive at a consensus score for the event that was then used in analyses. Cumulative adversity severity scores were computed by summing the maximum objective severity scores for each type of stressor endorsed (e.g., (15–17)).

After adding ELA as a covariate, the interaction between pollution burden and MPFC-DMN connectivity during *Negative Label > Negative Match* on the slope of depressive symptoms remained significant (Table S10).

Sensitivity Analyses: Examining the Intercept and Slope of Externalizing Symptoms as Outcomes

    Externalizing Symptoms. Participants completed the Youth Self Report (YSR; (18)), a self-report measure assessing emotional and behavioral problems. Participants rated 112 items on a 3-point scale (0=not true, 1=somewhat or sometimes, 2=very true or often true). We used the externalizing problems score, which comprise the aggressive behavior and rule-breaking behavior syndrome scales, in our analyses.

    As we did for depressive symptoms, we conducted linear mixed effects models with random intercepts and age slopes to estimate the intercepts and slopes of externalizing symptoms across the three timepoints for each participant. We modeled self-reported externalizing symptoms as a function of age (centered at 11 years, the mean age at T1) and extracted the intercept and slope parameter estimates for each participant, and tested whether pollution burden moderated associations between MPFC-DMN connectivity during *Negative Label > Negative Match* and both initial levels and slopes of externalizing symptoms.

    In contrast to our findings on depressive symptoms slope, pollution burden did not significantly moderate the association between MPFC-DMN connectivity *Negative Label > Negative Match* and slopes of externalizing symptoms (Table S11).

References

1. Dale AM, Fischl B, Sereno MI (1999): Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *Neuroimage* 9: 179–194.

2. Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Isik AI, Erramuzpe A, *et al.* (2019): fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods* 16. https://doi.org/10.1038/s41592-018-0235-4

3. Esteban O, Ciric R, Finc K, Blair RW, Markiewicz CJ, Moodie CA, *et al.* (2020): Analysis of task-based functional MRI data preprocessed with fMRIPrep. *Nat Protoc* 15: 2186–2202.

4. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL (2011): Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage* 54: 313–327.

5. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, *et al.* (2004): Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23: S208–S219.

6. Woolrich MW, Ripley BD, Brady M, Smith SM (2001): Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. *Neuroimage* 14: 1370–1386.

7. Behzadi Y, Restom K, Liau J, Liu TT (2007): A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* 37: 90–101.

8. Beckmann CF, Jenkinson M, Smith SM (2003): General multilevel linear modeling for group analysis in FMRI. *Neuroimage* 20: 1052–1063.

9. Beeney JE, Forbes EE, Hipwell AE, Nance M, Mattia A, Lawless JM, *et al.* (2021): Determining the key childhood and adolescent risk factors for future BPD symptoms using regularized regression: comparison to depression and conduct disorder. *J Child Psychol Psychiatry Allied Discip* 62: 223–231.

10. Verbeek JH, Whaley P, Morgan RL, Taylor KW, Rooney AA, Schwingshackl L, *et al.* (2021): An approach to quantifying the potential importance of residual confounding in systematic reviews of observational studies: A GRADE concept paper. *Environ Int* 157: 106868.

11. Fewell Z, Davey Smith G, Sterne JAC (2007): The Impact of Residual and Unmeasured Confounding in Epidemiologic Studies: A Simulation Study. *Am J Epidemiol* 166: 646–655.

12. Groenwold RHH, Klungel OH, Altman DG, van der Graaf Y, Hoes AW, Moons KGM (2013): Adjustment for continuous confounders: an example of how to prevent residual confounding. *C Can Med Assoc J = J l'Association medicale Can* 185: 401–406.

13. Adrian C, Hammen C (1993): Stress Exposure and Stress Generation in Children of Depressed Mothers. *J Consult Clin Psychol* 61: 354–359.

14. Rudolph KD, Hammen C (1999): Age and Gender as Determinants of Stress Exposure, Generation, and Reactions in Youngsters: A Transactional Perspective. *Child Dev* 70: 660–677.

15. Chahal R, Miller JG, Yuan JP, Buthmann JL, Gotlib IH (2022): An exploration of dimensions of early adversity and the development of functional brain network connectivity during adolescence: Implications for trajectories of internalizing symptoms. *Dev Psychopathol* 34: 557–571.

16. King LS, Graber MG, Colich NL, Gotlib IH (2020): Associations of waking cortisol with DHEA and testosterone across the pubertal transition: Effects of threat-related early life stress. *Psychoneuroendocrinology* 115: 104651.

17. Uy JP, Ho TC, Buthmann JL, Coury SM, Gotlib IH (2023): Early life stress, sleep disturbances, and depressive symptoms during adolescence: The role of the cingulum bundle. *Dev Cogn Neurosci* 63: 101303.

18. Achenbach TM, Rescorla LA (2001): Manual for the ASEBA School-Age Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Variables (Range)	Complete Data	Missing fMRI data	t or X^2	p
N	145	79		
Sex	87 females/58 males (60%/40%)	44 females/35 males (56%/44%)	0.16	.69
Age (T1) (9-13 years)	11.41 (1.06)	11.21 (1.01)	-1.38	.17
Race/Ethnicity			19.95	.0013
African-American	8 (5.5%)	11 (13.9%)		
Asian	19 (13.1%)	5 (6.33%)		
Biracial	29 (20%)	18 (22.8%)		
Hispanic	9 (6.2%)	11 (13.9%)		
Other	5 (3.45%)	9 (11.4%)		
White	75 (51.72%)	25 (31.6%)		
Pollution Burden Percentile (0.16-91.45)	29.80 (20.59)	38.57 (25.05)	-2.77	.006
Parent Education (2-8)	4.96 (1.24)	4.77 (1.40)	-1.02	.31
Income-to-Needs Ratio (0.05-1.97)	1.33 (0.52)	1.17 (0.62)	-1.87	.06
Depressive Symptoms (T1) (0-11)	2.23 (2.57)	2.01 (1.99)	0.67	.51

Table S1. Comparison of demographic and study variables between participants with at least one high-quality fMRI run and complete data (N=145) and those with missing/low-quality fMRI data (N-79).

Indicator (units)	CalEnviroScreen 3.0		CalEnviroScreen 4.0		p
	Measurement Time	Mean (SD)	Measurement Time	Mean (SD)	
ozone (ppm)	2012-2014	0.035 (0.004)	2017-2019	0.037 (0.004)	< .001
PM 2.5 (ug/m^3)	2012-2014	9.31 (0.94)	2015-2017	8.37 (0.51)	< .001
diesel PM (kg/day in 3.0; tons/year in 4.0)	2012	19.92 (12.15)	2016	0.20 (0.18)	--*
drinking water contaminants (contaminant index)	2005-2013	319.72 (210.71)	2011-2019	323.4 (111.54)	.62
pesticides (pounds/mile^2)	2012-2014	20.67 (173.21)	2017-2019	5.3 (35.01)	.14
toxic releases from facilities (toxicity-weighted concentration index)	2011-2013	272.61 (409.49)	2017-2019	317.26 (519.68)	.13
traffic density (vehicle-km per hour/km)	2013	902.79 (560.44)	2017	1259.17 (834.34)	< .001
cleanup sites (weighted distance)	2016	11.3 (21.57)	2021	10.36 (19.65)	.026
groundwater threats (weighted distance)	2016	26.96 (36.14)	2021	29.78 (38.84)	.008
hazardous waste facilities (weighted distance)	2012-2014	0.46 (1.56)	2018-2020	0.79 (1.58)	< .001
impaired water bodies (sum of number of pollutants found within water body)	2012	2.65 (3.59)	2018	2.75 (3.61)	.19
solid waste facilities (weighted distance)	2016	1.33 (3.98)	2021	1.13 (3.5)	.098
pollution burden score		34.6 (10.08)		34.86 (10.5)	.46

Table S2. Comparison of CalEnviroScreen indicators between versions 3.0 and 4.0. *Statistical

test not conducted due to difference in units reported.

	MPFC Activation (Negative Label > Negative Match)		
Predictors	**b**	**95% CI**	**p**
Intercept	-38.64	-94.96 – 17.68	0.178
Sex	3.14	-5.34 – 11.61	0.467
Age	1.88	-2.01 – 5.77	0.342
Household Socioeconomic Disadvantage	-4.49	-9.61 – 0.62	0.085
Neighborhood Socioeconomic Disadvantage	4.57	-1.12 – 10.25	0.115
Mean FD	53.7	16.46 – 90.94	0.005
Time Difference	-0.001	-0.014 – 0.012	0.893
Pollution Burden	0.4	0.18 – 0.61	<0.001
Hemisphere	-2.11	-8.98 – 4.76	0.546
Pollution Burden x Hemisphere	0.07	-0.12 – 0.26	0.474
Random Effects			
σ^2	284.05		
$\tau00_{ID}$	367.3		
ICC	0.56		
N_{ID}	143		
Observations	286		
Marginal R^2 / Conditional R^2	0.173 / 0.639		

Table S3. Results from linear mixed effects analyses testing whether the slope of pollution burden and MPFC activation during Negative Label > Negative Match differed between left and right hemispheres. FD = framewise displacement.

Region	Hemisphere	Voxels	x	y	z	Z max	p
Negative Match > Negative Label							
Occipital cortex/V1	L	27697	-8	-86	0	12.30	< .001
	R		12	-86	4	11.90	
Precentral gyrus	R	2714	46	6	26	7.94	< .001
	L	285	-28	-6	48	4.76	< .001
Amygdala	R	167	28	-2	-16	5.09	.004
Middle frontal gyrus	L	117	-44	4	34	5.22	.024
Positive Match > Positive Label							
Primary Visual/V1	L	23864	-8	-86	0	11.6	< .001
	R		14	-84	2	11.50	
Precentral gyrus	R	583	48	6	26	6.37	< .001
Superior frontal gyrus	R	546	22	-2	50	6.43	< .001
Intraparietal sulcus	L	167	-44	-38	48	4.98	.005
Thalamus	L	160	-22	-28	-2	5.97	.007
Amygdala	R	142	24	0	-16	5.05	.012
VMPFC	L	105	-8	62	-4	4.90	.045

Table S4. Results from whole-brain analyses identifying regions that were more active during

Match relative to Label conditions for Negatively- vs. Positively-valenced stimuli, cluster-

corrected at $Z > 3.1$, $p < .05$.

Region	Hemisphere	Voxels	x	y	z	Z max	p
Negative Label > Negative Match							
Middle temporal gyrus	L	7367	-44	-40	2	6.96	< .001
Posterior cingulate	L	2548	-2	-24	38	4.94	< .001
Central opercular cortex	R	175	38	6	14	3.98	.003
Superior frontal gyrus	L	139	-8	14	60	4.96	.011
Cerebellum	R	123	28	-72	-32	4.74	.020
Precuneus	L	118	-22	-48	32	5.02	.024
Positive Label > Positive Match							
Superior temporal gyrus	L	3183	-52	-34	0	9.19	< .001
Middle temporal gyrus	R	849	50	-30	-6	5.96	< .001
Superior frontal gyrus	L	2037	-2	32	38	5.77	< .001
Inferior frontal gyrus	L	1963	-52	20	24	7.41	< .001
Frontal pole	L	1541	-22	52	32	5.73	< .001
	R	1249	32	60	6	5.63	< .001
Angular gyrus	R	969	54	-54	46	5.92	< .001
DLPFC	L	539	-40	6	46	4.89	< .001
	R	440	44	24	38	5.22	< .001
Orbitofrontal/Insula cortex	R	491	48	22	-10	4.86	< .001
Cerebellum	R	317	26	-72	-28	5.46	< .001
Cingulate gyrus	L	311	-4	-10	34	4.50	< .001
Temporal Pole	L	202	-50	10	-20	4.38	.002
(Positive Label > Positive Match) > (Negative Label > Negative Match)							
Paracingulate gyrus	R	230	6	24	52	4.07	< .001
DLPFC	R	182	44	36	40	4.12	.004

Table S5. Results from whole-brain analyses identifying regions that were more active during

Label relative to Match conditions for Negatively- vs. Positively-valenced stimuli, and regions

that showed a difference in Label > Match by valence, cluster-corrected at Z > 3.1, p < .05.

	Depressive Symptoms							
	Baseline (Intercept)				**Longitudinal Changes (Slope)**			
	b	**SE**	**t**	**p**	**b**	**SE**	**t**	**p**
Intercept	1.19	0.48	2.49	.01	0.24	0.20	1.21	.23
Sex	0.12	.20	0.62	.54	0.10	0.08	1.34	.18
Age	-0.02	0.09	-0.26	.80	-0.05	0.03	-1.53	.13
Household Socioeconomic Disadvantage	0.10	0.11	0.90	.37	0.01	0.05	0.29	.77
Neighborhood Socioeconomic Disadvantage	0.02	0.13	0.13	.90	-0.01	0.05	-0.26	.80
Pollution Burden Percentile	-0.004	0.005	-0.85	.40	-0.001	0.002	-0.63	.53
Mean FD	-0.16	0.84	-0.20	.84	-0.0004	0.33	-0.001	.99
Time Difference	0.001	0.0003	2.26	.03	-0.000004	0.0001	-0.03	.97
MPFC-IFG Connectivity	0.012	0.10	0.12	.90	-0.05	0.04	-1.20	.23
MPFC-IFG Connectivity x Pollution Burden Percentile	-0.0005	0.005	-0.09	.93	-0.0004	0.002	-0.19	.85
Baseline Depressive Symptoms	--	--	--	--	0.25	0.05	5.31	<.001
Baseline Depressive Symptoms (Quadratic)	--	--	--	--	-0.14	0.03	-4.75	<.001

Table S6. Results from regression analyses testing whether MPFC-IFG connectivity during

Negative Label > Negative Match was related to baseline (left) and/or longitudinal changes

(right) in depressive symptoms, and whether association differed by pollution burden. Analysis

examining longitudinal changes in depressive symptoms additional controlled for baseline

depressive symptoms (linear and quadratic). Age, baseline depressive symptoms, and pollution

burden percentile are mean-centered.

Variable	Coefficients
Intercept	0.3118
Age	
Sex	0.0349
Neighborhood Socioeconomic Disadvantage	
Household Socioeconomic Disadvantage	
Time Difference	0.0138
Mean FD	-0.0240
Depressive Symptoms Intercept	0.1491
Depressive Symptoms Intercept (Quadratic)	-0.0690
MPFC-DMN Connectivity	0.1022
ozone	
PM 2.5	
diesel PM	0.0402
drinking water	-0.0087
pesticides	-0.0408
toxic releases	
traffic density	-0.0007
cleanup sites	
groundwater threats	
hazardous waste facilities	
impaired water bodies	0.0624
solid waste facilities	-0.0906
MPFC-DMN Connectivity x ozone	0.1724
MPFC-DMN Connectivity x PM 2.5	
MPFC-DMN Connectivity x diesel PM	0.0444
MPFC-DMN Connectivity x drinking water	
MPFC-DMN Connectivity x pesticides	
MPFC-DMN Connectivity x toxic releases	0.0622
MPFC-DMN Connectivity x traffic density	
MPFC-DMN Connectivity x cleanup sites	0.0404
MPFC-DMN Connectivity x groundwater threats	0.0209
MPFC-DMN Connectivity x hazardous waste facilities	0.0514
MPFC-DMN Connectivity x impaired water bodies	-0.1261
MPFC-DMN Connectivity x solid waste facilities	

Table S7. Results from LASSO regression analysis predicting slopes of depressive symptoms.

	Depressive Symptoms Slope			
	b	SE	t	p
Intercept	0.23	0.2	1.22	.22
Sex	0.12	0.07	1.22	.18
Age	-0.05	0.03	-1.34	.11
Household Socioeconomic Disadvantage	0.02	0.05	0.36	.72
Neighborhood Socioeconomic Disadvantage	-0.01	0.05	-0.24	.81
Mean FD	-0.04	0.3	-0.13	.90
Time Difference	-0.0001	0.0001	-0.45	.65
Pollution Burden Percentile	-0.002	0.002	-1.12	.27
MPFC-DMN Connectivity	0.06	0.05	1.15	.25
MPFC-DMN Connectivity x Pollution Burden Percentile	0.006	0.002	2.40	.018
Baseline Depressive Symptoms	0.25	0.05	5.38	< .001
Baseline Depressive Symptoms (Quadratic)	-0.13	0.03	-4.76	< .001

Table S8. Results from regression analyses testing whether MPFC-DMN connectivity during

Negative Label > Negative Match was related to longitudinal changes in depressive symptoms,

and whether association differed by pollution burden, additionally controlling for baseline

depressive symptoms (linear and quadratic). Age, baseline depressive symptoms, and pollution

burden percentile are mean-centered.

	Depressive Symptoms Slope		
	b	95% CI	p
Intercept	0.24	-0.14 – 0.62	0.217
Age	-0.05	-0.11 – 0.02	0.183
Sex	0.12	-0.03 – 0.26	0.117
Household Socioeconomic Disadvantage	0.02	-0.07 – 0.11	0.716
Neighborhood Socioeconomic Disadvantage	-0.01	-0.11 – 0.09	0.812
Pollution Burden Percentile	-0.002	-0.006 – 0.002	0.256
MPFC-DMN Connectivity	0.06	-0.04 – 0.17	0.255
Time Difference	-0.0001	-0.0003 – 0.0002	0.642
Mean FD	-0.04	-0.68 – 0.60	0.891
Depressive Symptoms Baseline	0.25	0.16 – 0.34	<0.001
Depressive Symptoms Baseline (Quadratic)	-0.13	-0.19 – -0.08	<0.001
MPFC-DMN Connectivity x Pollution Burden Percentile	0.01	0.0012 – 0.01	0.018
Random Effects			
σ^2	0.14		
$\tau_{00 \ census \ tract}$	0.01		
ICC	0.04		
$N_{census \ tract}$	125		
Observations	145		
Marginal R^2 / Conditional R^2	0.251 / 0.283		

Table S9. Results from multilevel regression analyses using census tract as a grouping variable testing whether MPFC-DMN connectivity during Negative Label > Negative Match was related to longitudinal changes in depressive symptoms, and whether association differed by pollution burden. Age, baseline depressive symptoms, and pollution burden percentile are mean-centered.

	Depressive Symptoms							
	Baseline (Intercept)				**Longitudinal Changes (Slope)**			
	b	**SE**	**t**	**p**	**b**	**SE**	**t**	**p**
Intercept	0.90	0.46	1.97	.05	0.28	0.20	1.40	.16
Sex	.19	0.18	1.05	.30	0.11	0.07	1.49	.14
Age	-0.02	0.08	-0.27	.79	-0.04	0.03	-1.31	.19
Household Socioeconomic Disadvantage	0.0004	0.11	0.03	.99	0.03	0.05	0.62	.53
Neighborhood Socioeconomic Disadvantage	.001	0.12	0.11	.99	-0.01	0.05	-0.26	.80
Mean FD	-0.37	0.79	-0.47	.64	-0.02	0.32	-0.07	.94
Time Difference	0.005	0.0003	1.85	.07	-0.00004	0.0001	-0.37	.71
Pollution Burden Percentile	-0.01	0.005	-1.14	.25	-0.002	0.002	-1.12	.27
MPFC-DMN Connectivity	0.17	0.13	1.30	.19	0.06	0.05	1.06	.29
MPFC-DMN Connectivity x Pollution Burden Percentile	-0.005	0.006	-0.77	.44	0.01	0.002	2.49	.014
Baseline Depressive Symptoms	--	--	--	--	0.26	0.05	5.46	<.001
Baseline Depressive Symptoms (Quadratic)	--	--	--	--	-0.13	0.03	-4.78	<.001
Early Life Adversity	0.06	0.016	3.59	<.001	-0.01	0.007	-0.97	.33

Table S10. The interaction between pollution burden percentile (mean-centered) and MPFC-DMN connectivity during *Negative Label > Negative Match* on slopes of depressive symptoms remained significant after additionally controlling for early life adversity.

	Externalizing Symptoms							
	Baseline (Intercept)				Longitudinal Changes (Slope)			
	b	SE	t	p	b	SE	t	p
Intercept	9.82	1.69	5.79	< .001	0.29	0.39	0.74	.46
Sex	-0.89	0.67	-1.33	.19	0.04	0.15	0.25	.80
Age	0.003	0.30	.012	.99	-0.03	0.07	-0.37	.71
Household Socioeconomic Disadvantage	0.79	0.43	1.86	.07	-0.02	0.10	-0.18	.85
Neighborhood Socioeconomic Disadvantage	-0.07	0.45	-0.16	.87	-0.10	0.10	-0.92	.36
Mean FD	3.26	2.098	1.09	.28	0.52	0.69	0.76	.45
Time Difference	0.0003	0.001	0.32	.75	-0.0002	0.0002	-0.69	.49
Pollution Burden Percentile	-0.01	0.018	-0.30	.76	-0.01	0.004	-1.74	.08
MPFC-DMN Connectivity	-0.18	0.49	-0.37	.71	0.15	0.11	1.37	.17
MPFC-DMN Connectivity x Pollution Burden Percentile	-0.02	0.02	-1.12	.27	0.01	0.005	1.76	.08
Baseline Externalizing Symptoms	--	--	--	--	-0.002	0.02	-0.12	.91
Early Life Adversity	0.29	0.06	4.56	< .001	-0.0001	0.016	-0.006	.99

Table S11. The interaction between pollution burden percentile (mean-centered) and MPFC-

DMN connectivity during *Negative Label > Negative Match* on externalizing symptoms

intercepts and slopes were not significant.
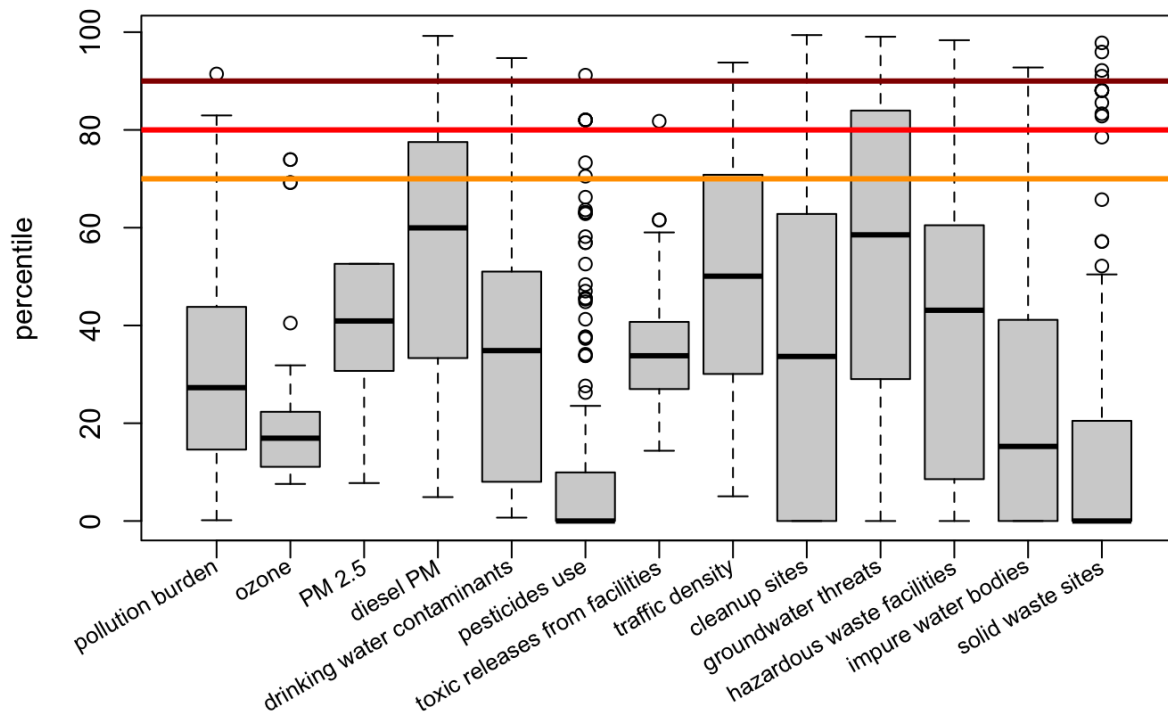
Figure S1. Boxplots representing the distribution of pollution burden percentiles and associated indicators. Orange, red, and dark red lines indicate 70th, 80th, and 90th percentiles, respectively.
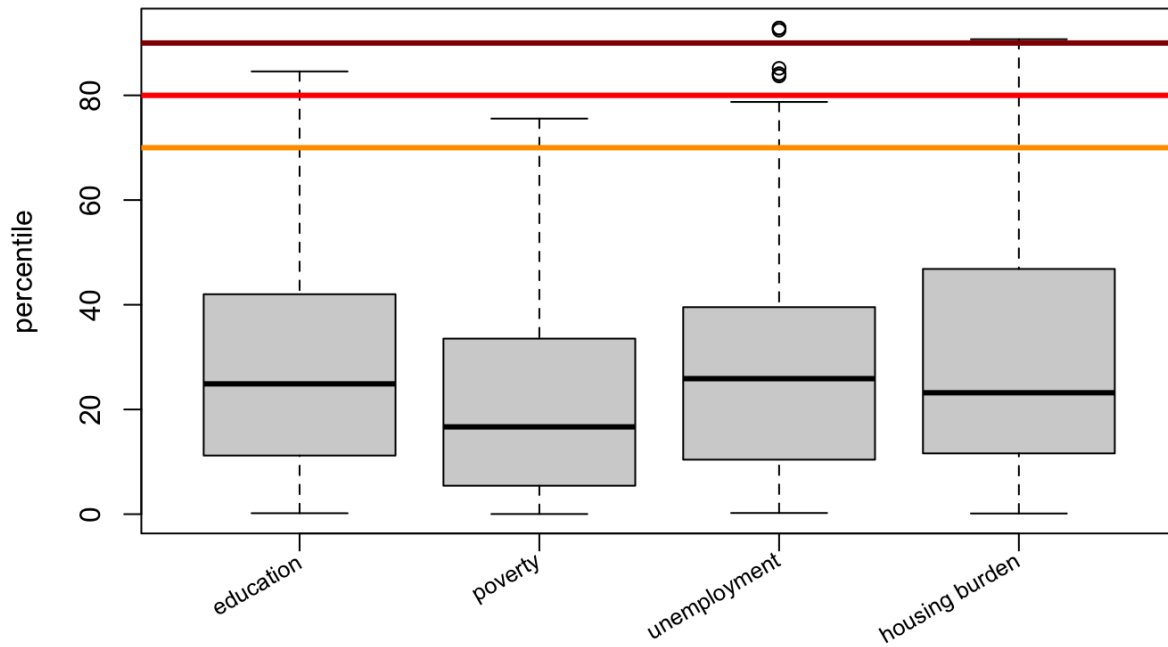
Figure S2. Boxplots representing the distribution of percentiles of neighborhood socioeconomic disadvantage indicators. Higher percentiles indicate higher levels of disadvantage. Orange, red, and dark red lines indicate 70th, 80th, and 90th percentiles, respectively.
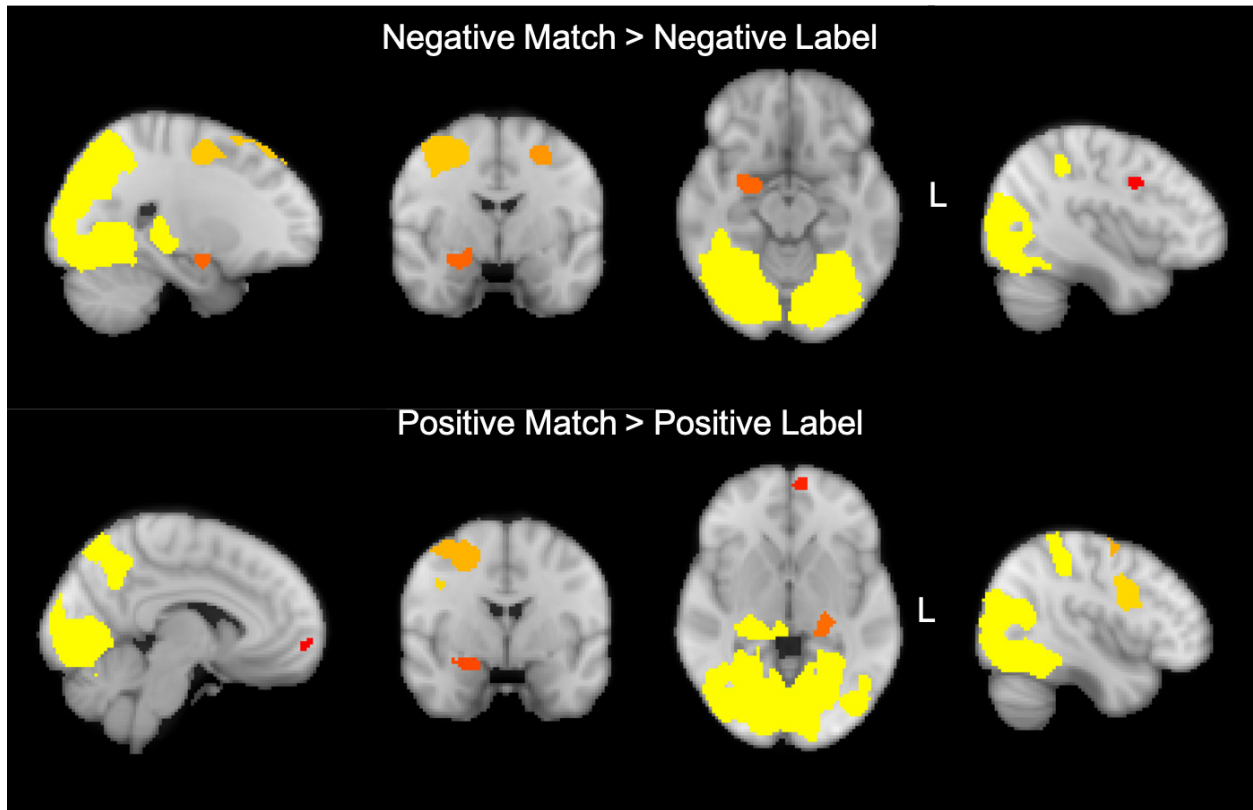
Figure S3. Results from whole-brain cluster analyses depicting clusters/regions that were more active during Negative Match > Negative Label (top) and during Positive Match > Positive Label (bottom), cluster-corrected at Z > 3.1, p < .05.
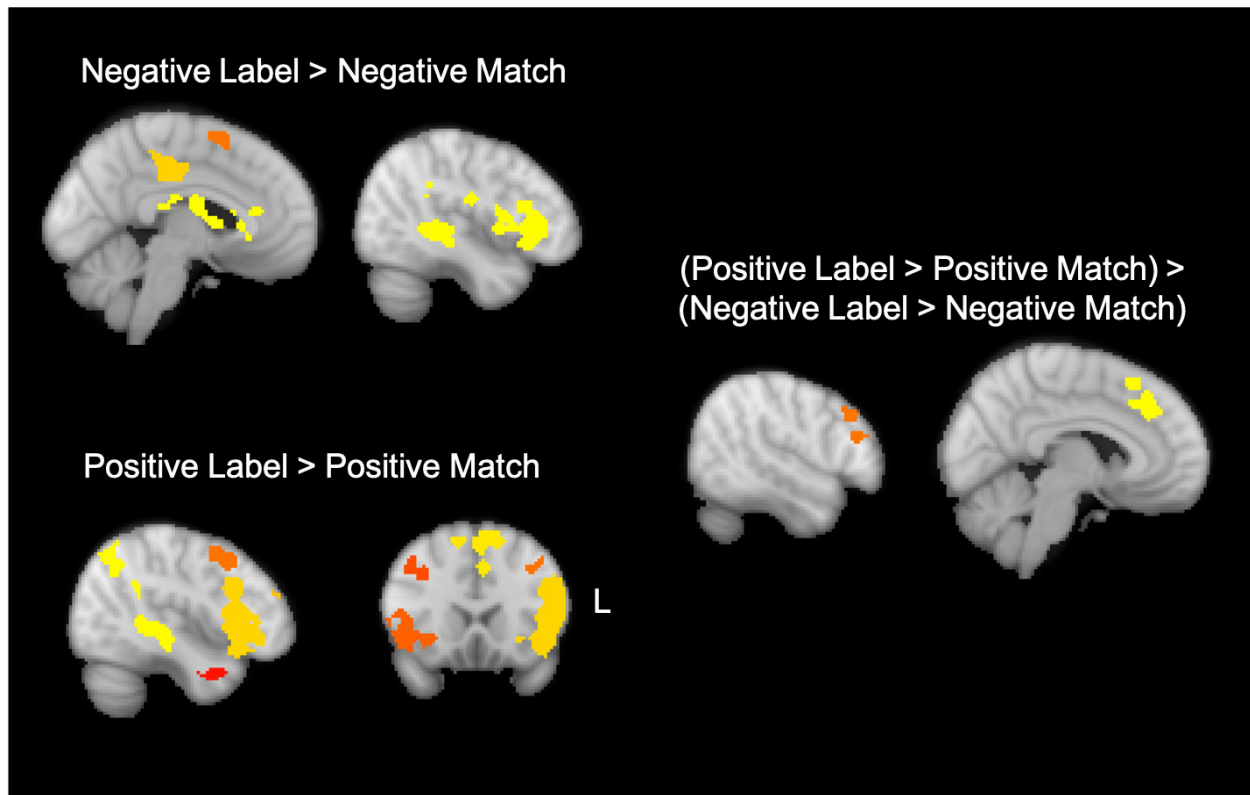
Figure S4. Results from whole-brain cluster analyses depicting clusters/regions that were more active during Negative Label > Negative Match (top left) and during Positive Label > Positive Match (bottom left), and regions that differed between Positive and Negative valence, cluster-corrected at Z > 3.1, p < .05.
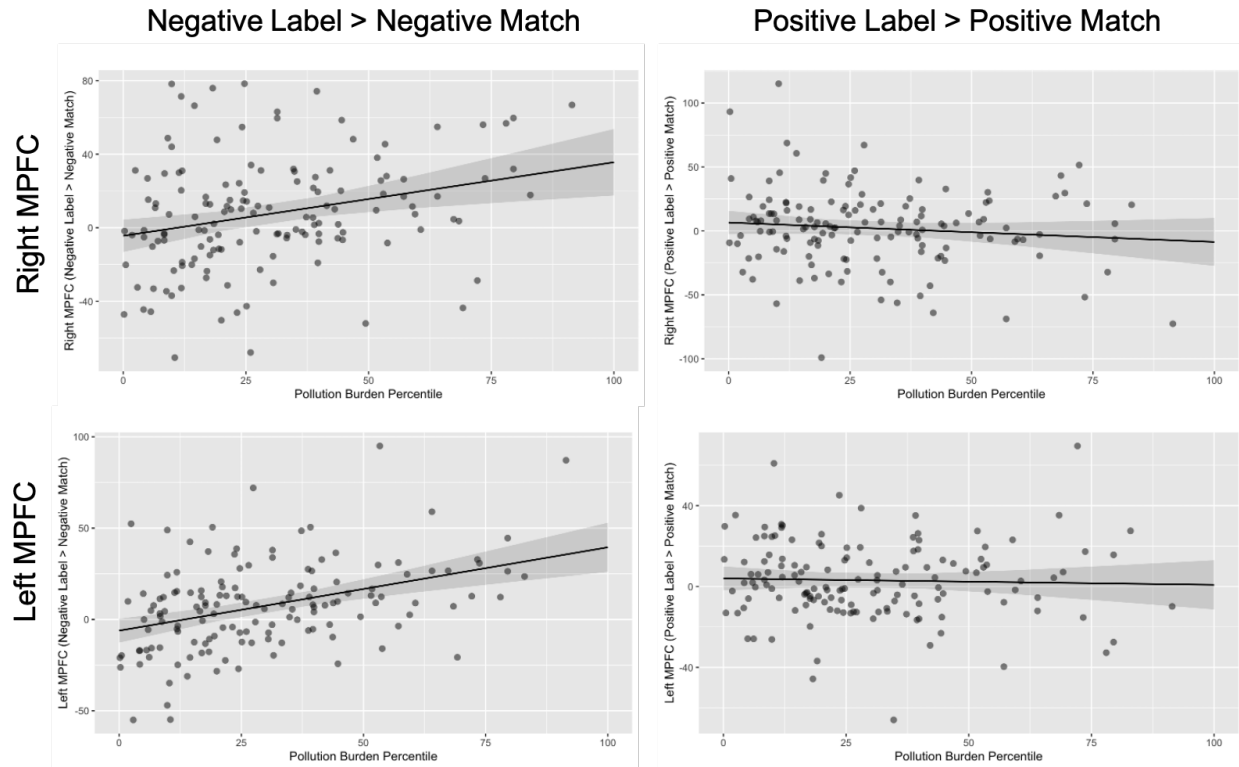
Figure S5. Scatterplots depicting associations of pollution burden percentile with activation in right MPFC (top) and left MPFC (bottom) during Negative Label > Negative Match (left) and Positive Label > Positive Match (right).
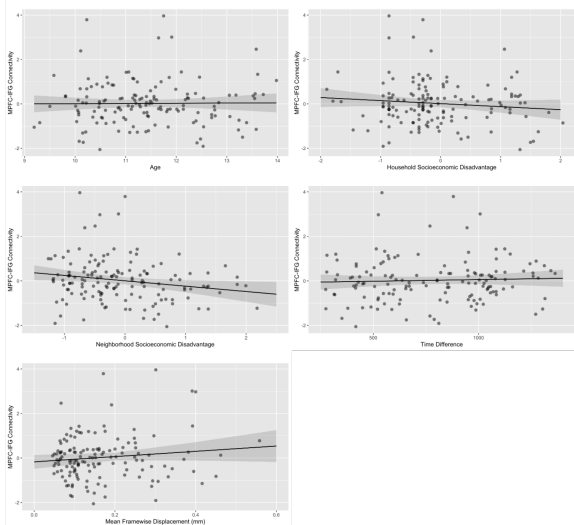
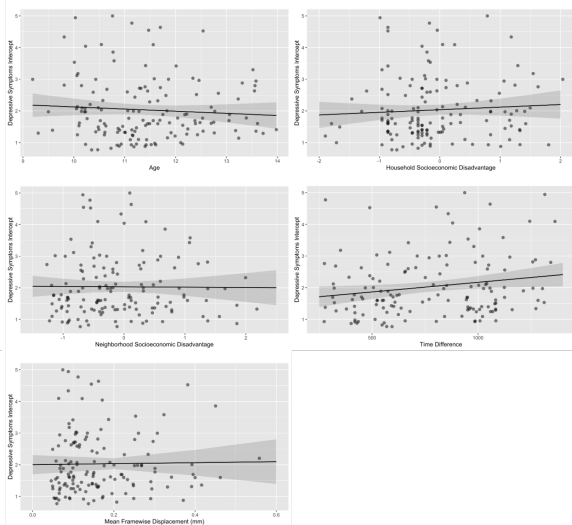**A) Bivariate Plots of Bilateral MPFC Activation and Covariates**
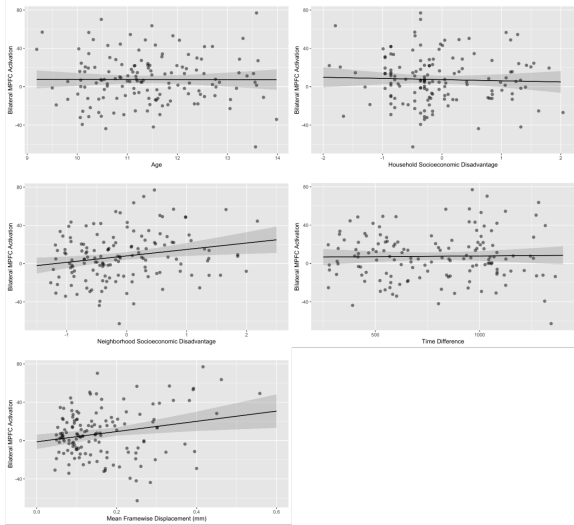
B) Bivariate Plots of MPFC-DMN Connectivity and Covariates
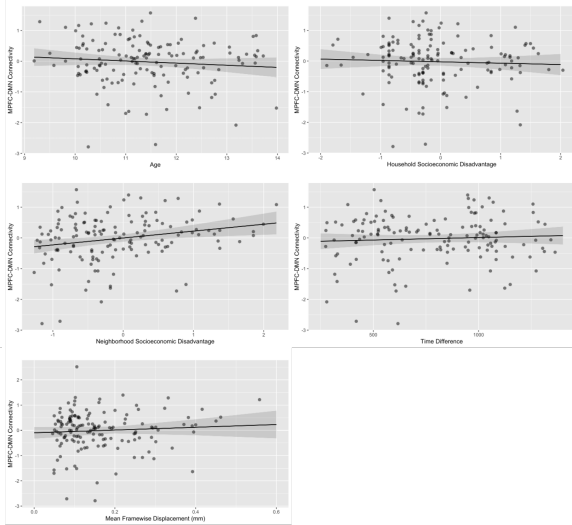
C) Bivariate Plots of MPFC-IFG Connectivity and Covariates

D) Bivariate Plots of Depressive Symptoms Intercepts and Covariates

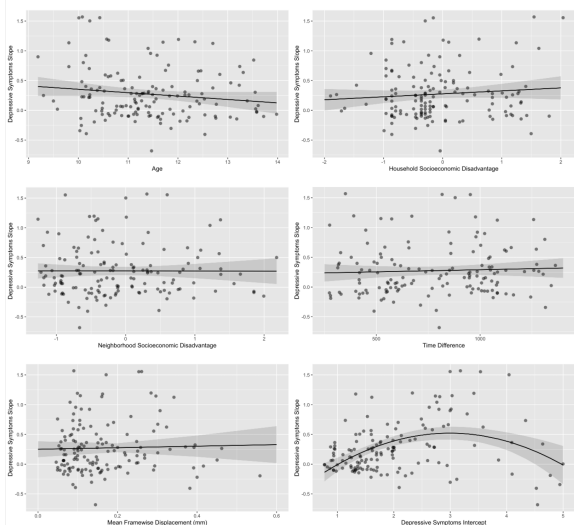**E) Bivariate Plots of Depressive Symptoms Slope and Covariates**

Figure S6. Bivariate plots of covariates (age at scan, household socioeconomic disadvantage, neighborhood socioeconomic disadvantage, time difference between pollution measurement and brain scan, and mean framewise displacement) with each outcome variable - (A) bilateral MPFC activation, (B) MPFC-DMN connectivity, (C) MPFC-IFG connectivity, (D) depressive symptoms intercept. For (E) depressive symptoms slopes, bivariate plot with depressive symptoms intercept is also included.