Correction

NEUROSCIENCE, PSYCHOLOGICAL AND COGNITIVE SCIENCES

Correction for "Patterns of sociocognitive stratification and perinatal risk in the child brain," by Dag Alnæs, Tobias Kaufmann, Andre F. Marquand, Stephen M. Smith, and Lars T. Westlye, which was first published May 14, 2020; 10.1073/pnas.2001517117 (*Proc. Natl. Acad. Sci. U.S.A.* **117**, 12419–12427).

The authors wish to note the following: "We discovered that for a subset of the behavioral measures, ∼20% of the participants had the 1-y follow-up data included instead of the baseline measures. Reanalysis using baseline data only yielded close to identical results, with mode subject weights as well as mode variable weights correlating $r > .98$ between the original and corrected analysis. In summary, all results are retained, and none of the conclusions change. A PDF showing the original and updated results as well as figures is available in the OSF repository associated with the manuscript: https://osf.io/m39yk/."

Published August 30, 2021.

www.pnas.org/cgi/doi/10.1073/pnas.2113760118

CORRECTION

Patterns of sociocognitive stratification and perinatal risk in the child brain

Dag Alnæs[a,1], Tobias Kaufmann[a], Andre F. Marquand[b,c,d], Stephen M. Smith[e], and Lars T. Westlye[a,f,g]

[a]Norwegian Centre for Mental Disorders Research (NORMENT), Institute of Clinical Medicine, University of Oslo & Division of Mental Health and Addiction, Oslo University Hospital, 0424 Oslo, Norway; [b]Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen 6525EN, The Netherlands; [c]Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen 6525EN, The Netherlands; [d]Department of Neuroimaging, Centre for Neuroimaging Sciences, Institute of Psychiatry, King's College London, SE5 8AF London, United Kingdom; [e]Wellcome Centre for Integrative Neuroimaging (WIN FMRIB), Oxford University, OX3 9DU Oxford, United Kingdom; [f]KG Jebsen Centre for Neurodevelopmental Disorders, Institute of Clinical Medicine, University of Oslo, 0318 Oslo, Norway; and [g]Department of Psychology, University of Oslo, 0373 Oslo, Norway

Edited by Marcus E. Raichle, Washington University in St. Louis, St. Louis, MO, and approved April 8, 2020 (received for review January 26, 2020)

The expanding behavioral repertoire of the developing brain during childhood and adolescence is shaped by complex brain–environment interactions and flavored by unique life experiences. The transition into young adulthood offers opportunities for adaptation and growth but also increased susceptibility to environmental perturbations, such as the characteristics of social relationships, family environment, quality of schools and activities, financial security, urbanization and pollution, drugs, cultural practices, and values, that all act in concert with our genetic architecture and biology. Our multivariate brain–behavior mapping in 7,577 children aged 9 to 11 y across 585 brain imaging phenotypes and 617 cognitive, behavioral, psychosocial, and socioeconomic measures revealed three population modes of brain covariation, which were robust as assessed by cross-validation and permutation testing, taking into account siblings and twins, identified using genetic data. The first mode revealed traces of perinatal complications, including preterm and twin birth, eclampsia and toxemia, shorter period of breastfeeding, and lower cognitive scores, with higher cortical thickness and lower cortical areas and volumes. The second mode reflected a pattern of sociocognitive stratification, linking lower cognitive ability and socioeconomic status to lower cortical thickness, area, and volumes. The third mode captured a pattern related to urbanicity, with particulate matter pollution (PM^{25}) inversely related to home value, walkability, and population density, associated with diffusion properties of white matter tracts. These results underscore the importance of a multidimensional and interdisciplinary understanding, integrating social, psychological, and biological sciences, to map the constituents of healthy development and to identify factors that may precede maladjustment and mental illness.

population imaging | neurodevelopment | childhood/adolescence | psychology | neuroscience

The complexity and idiosyncratic characteristics of the human mind originate in an intricate web of interactions between genes; brain circuits; behaviors; and economic, social, and cultural factors during childhood and adolescence. The major life changes associated with the transition into young adulthood offer opportunities for adaptation and growth but also increased susceptibility to detrimental perturbations, such as the characteristics of social and parental relationships, family environment, quality of schools and activities, economic security, urbanization and pollution, drugs, cultural practices, and values, that all act in concert with our genetic architecture and biology (1). A multidimensional understanding of the interplay of these factors is paramount to identify the constituents of healthy development and to identify factors that may precede maladjustment and mental illness.

Population-based neuroimaging now allows us to take a bird's-eye view on this stupendous multiplicity and to bring hitherto unseen patterns into focus (2). The Adolescent Brain Cognitive Development (ABCD) study (3) currently provides brain images of more than 10,000 children aged 9 to 11 y at baseline. The participants are recruited and assessed across 21 sites in the United States, making it one of the largest neurodevelopmental imaging resources available. In addition to multimodal MRI, the ABCD study includes a broad range of cognitive, behavioral, clinical, psychosocial, and socioeconomic measures. Importantly, these participants will be followed up for over a 10-y period, allowing the opportunity to reveal robust population-level multivariate brain–behavior associations in neurodevelopment and to later assess their predictive value as follow-up data become available. While each neuroimaging feature typically explains a minute amount of unique variance in behavioral outcome (4, 5), their combined predictive value is nonnegligible, including predictive patterns for identification of individuals (6, 7) and characteristics such as age (8, 9), cognitive ability (10), and psychopathology (10). This added value of multivariate and combinatorial approaches for prediction of complex traits is highly analogous to the substantial polygenic accumulation of small

Significance

Adolescence is a transition period between childhood to adulthood, associated with heightened sensitivity to the social and cultural environment, and offers opportunities for adaptation and growth but also increased susceptibility to detrimental perturbations. We reveal three population-level patterns in children aged 9 to 11, linking measures of brain morphology and white matter fiber properties to cognitive, behavioral, psychosocial, and socioeconomic measures. They reveal associations with early life factors, such as preterm birth, birth complications, and duration of breastfeeding; socioeconomic and cognitive stratification; and urbanicity, home value, and air pollution. The results support that political priorities and decisions aiming to improve health outcomes and adaptation during transformative life phases should be based on interdisciplinary perspectives integrating social, psychological, and biological sciences.

Author contributions: D.A., T.K., A.F.M., S.M.S., and L.T.W. analyzed data and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: Tables containing the included variables, and code implementing the analysis, are publicly available at the following Open Science Framework (OSF) repository: https://doi.org/10.17605/OSF.IO/M39YK.

[1]To whom correspondence may be addressed. Email: dag.alnas@psykologi.uio.no.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2001517117/-/DCSupplemental.

First published May 14, 2020.

effects in the genetic architecture of complex human traits and disorders (11, 12).

Adolescence is a transition period between childhood and adulthood and a period of protracted brain maturation, associated with heightened sensitivity to the social and cultural environment (13). For most individuals, this transition results in successful acquirement of skills and coping strategies required for adulthood and subsequent independence from caregivers; however, it is also a period of increased risk for mental health issues (14). Vulnerability is associated with neurodevelopmental factors preceding illness, including growing up in low-socioeconomic (SES) homes and childhood deprivation as well as reduced brain integrity already in early life (15). Mapping positive and negative factors impacting the brain as well as psychological adjustment before and during the transition from childhood to adulthood is therefore of pivotal importance. Combining levels of information using latent-variable approaches which model all available information may reveal interpretable patterns among multiple brain imaging features and variables such as cognition and sociodemographics (5, 16). One recent example revealed that a wide range of cognitive, clinical, and lifestyle measures constitute a positive–negative dimension associated with adult brain network functional connectivity (2).

Here we used an analogous approach in 7,577 children aged 9 to 11 from the ABCD study, collected across 21 sites across the United States, combining canonical correlation analysis (CCA) with independent component analysis (ICA) to derive population-level modes of covariation, linking behavioral, psychosocial, socioeconomic, and demographical variables (behavioral measures) to a wide set of neuroimaging phenotypes. Each resulting mode represents an association between a linear combination of behavioral measures with a separate combination of imaging features that show similar variation across participants (16). In order to avoid overfitting, which is particularly important when employing data-driven approaches, and due to the high number of intercorrelated features, CCA was performed after data reduction with principal component analysis. Robustness and reliability of the identified modes were assessed using stratified cross-validation to avoid estimating the modes using one individual in a sibling/twin pair and deriving out-of-sample scores for the other, and permutation testing with restricted exchangeability, taking into account siblings and twins based on participants' genetic data. To express results in the original variable space, CCA–ICA subject weights were correlated back into the original data. Based on earlier reports of population-level associations between measures of life outcomes and brain connectivity (2) and structure (17) in adults and the known and rising socioeconomic inequalities in the United States (18), as well as the impact of socioeconomic factors on child brain development (19, 20), we expected to find traces of social stratification in the child brain.

Methods

ABCD Data Access. We accessed baseline MRI, behavioral, clinical, and genetic data from ABCD Annual curated release 2.0.1. A flowchart describing the processing and analysis steps is provided in *SI Appendix*, Fig. S1.

Behavioral, Clinical, Cognitive, and Demographical Data. Tabulated data were imported and processed using R (https://cran.r-project.org). We accessed data from 11,853 participants. *SI Appendix*, Table S1, lists the behavioral measures included the analysis. We used the function nearZeroVar from the R package caret (v. 6.0-81, https://github.com/topepo/caret/) to identify and exclude any continuous variables with zero or near-zero variance and categorical variables with a ratio of >0.95 for the most common compared to the second most common response. For each remaining variable we derived robust z scores by calculating each score's absolute deviation from the median absolute deviation (MAD) (21) and removed values with a $z > 4$ ($4 \times$ MAD). Those with a $z > 3$ were manually inspected: e.g., measures of facility income, time spent on phone, and several measures of area deprivation have scores with $z > 3$ but were kept in the analysis. We then excluded variables with less than 90% of retained datapoints, before excluding subjects with less than 90% retained data across the retained variables. The remaining subjects ($n = 11{,}809$) were included for further analysis.

MRI Imaging-Derived Phenotypes. We accessed T1-weighted and T2-weighted ($n = 11{,}534$) and diffusion-weighted (DWI, $n = 11{,}400$) tabulated MRI data from ABCD curated release 2.0.1. *SI Appendix*, Table S2, lists the MRI features included in the analysis. We included participants who passed quality assurance using the recommended quality control (QC) parameters (T1, $n = 11{,}359$; T2, $n = 10{,}476$; DWI, $n = 10{,}414$) described in the ABCD 2.0.1 Imaging Instruments Release Notes and who had all included modalities available ($n = 9{,}811$). ABCD preprocessing and QC steps are described in detail in the methodological reference for the ABCD study (22). For each included imaging phenotype, we calculated the MAD for each score and removed values with MAD >3. Subjects with less than 90% of features retained in any of the imaging modalities and features with less than 90% of retained subjects were excluded from analysis. The remaining subjects ($n = 9{,}016$) were included for further analysis.

Genetic Data. We accessed genetic data for 10,627 participants to identify siblings and twins. We used genome-wide complex trait analysis (23) to create a genetic relationship matrix after performing the following filtering: removal of SNPs in the major histocompatibility complex (25:35 Mb region on chr6) and the inversion region of chr8 (7:13 Mb) and SNPs with genotyping rate <99%, minor allele frequency <5%, and pairwise pruning of SNPs in linkage disequilibrium ($r^2 > 0.2$, window of 5,000, step of 500). To account for siblings and twins in the dataset, three groups were created based on the following genetic relatedness cutoffs: nonsiblings, <0.4; siblings, >0.4 and <0.6; and monozygotic (MZ) twins, >0.8, with the two latter groups containing pairs of siblings/MZ twins. These three groups were used for stratified cross-validation and creation of permutation exchangeability blocks.

CCA. We performed CCA (24) using MATLAB R2019b. Participants with MRI, behavioral, and genetic data ($n = 7{,}577$; mean age = 9.9 y, SD = 0.63 y, males = 3,957) were included. We applied a rank-based normal transformation to the behavioral/clinical data using palm_inormal from the FMRIB Software Library Permutation Analysis of Linear Models (25) (v. 0.52, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/PALM). Next, we residualized all measures with respect to age and sex using linear models. Imaging phenotypes were also residualized for site/scanner (using the scanners' unique serial numbers), and volumetric features were also corrected for estimated total intracranial volume (eTIV from Freesurfer). eTIV was also included as variable in the analysis to capture associations with global volume, in addition to the eTIV-corrected volumes capturing associations with regional specificity. For both MRI and behavioral measures, missing values were imputed with knnimpute, replacing missing data based on the k nearest-neighbor columns based on Euclidian distance ($k = 3$). An alternative approach without imputation is described below and did not change results. Data were then z-normalized and submitted (separately for imaging and behavioral data) to principal component analysis (PCA; *SI Appendix*, Fig. S2), to avoid issues with rank deficiency and to increase robustness of estimated modes by avoiding fitting to noise. We extracted the first 200 components for both the imaging and behavioral data and submitted these to CCA.

Cross-Validation. To assess the reliability and generalizability of the resulting CCA modes we performed the following 10-fold cross-validation procedure: For each iteration ($n = 100$) of the cross-validation loop the dataset was randomly divided into 10 folds, stratified by the genetic relatedness groups, and ensuring that sibling and twin pairs were kept together to avoid training on one sibling/twin in a pair and testing on the other. While keeping each fold (10% of participants) out once, we submitted the remaining data (90% of participants) to PCA (separately for imaging and MRI data) and then to CCA. Next, we multiplied the kept-out behavioral measure and MRI feature matrices with the estimated PCA coefficient matrices, before multiplying the resulting PCA scores with the canonical coefficients and then correlated the resulting CCA scores. Finally, we took the average of these canonical correlations across the 10 folds (*SI Appendix*, Fig. S3). This procedure was repeated 100 times to derive mean canonical correlations for kept-out data and used for calculating P values after permutation testing. We also correlated the CCA subject measure and MRI coefficients derived for kept-out participants with those from the full analysis.

Permutation Testing. To assess significance of the resulting CCA modes, we ran 1,000 iterations of the same 10-fold cross-validation procedure described above but with the order of participants of the imaging phenotype matrix randomly permuted in each iteration, respecting twin/sibling relationships and collecting canonical correlations for the kept rather than the kept-out data to account for overfitting by the CCA. We then collected the maximum canonical correlation across CCA modes (i.e., mode 1) for each permutation to form a null distribution to calculate familywise error corrected *P* values. *P* values for each of the CCA modes were calculated by dividing the count of permuted maximum R values (including the observed value) ≥ the mean of cross-validated R values by the number of permutations. CCA modes with a corrected *P* value <0.01 were included for further analysis (*SI Appendix*, Fig. S3).

CCA–ICA. The canonical variates become increasingly difficult to interpret due to their orthogonality. Since we had more than one significant mode, and following procedures described by Miller et al. (16), we used ICA to obtain more interpretable modes: we extracted and combined the behavioral and MRI CCA scores for the three significant variates, correlated these with the original data matrix, transformed the correlations using a Fisher Z-transform, and submitted these to ICA. We estimated three components (the number of significant and extracted CCA canonical pairs) using fastICA (26). The ICA subject weight correlations with the canonical variates are shown in *SI Appendix*, Fig. S4. To assess the reliability and generalizability of the ICA decomposition we reran 100 iterations of the 10-fold cross-validation procedure described above, this time including ICA estimation after the PCA and CCA step, and then correlated ICA subject weights derived from kept-out data to those from the full analysis (*SI Appendix*, Fig. S3). To assess and plot the significant CCA–ICA modes in the full original variable space, we correlated the subject weights for each CCA–ICA mode with the original age- and sex (+eTIV)-adjusted matrices. For each significant mode of population covariation, we also plotted the variable text/descriptions for the 35 variables with the highest explained variance in the original adjusted data (lists of all variables and associated descriptions, correlations, and ICA weights can be found in *SI Appendix*, Tables S3–S5). ICA subject weight histograms are shown in *SI Appendix*, Fig. S5. The explained variance of single variables ranged between 10 and 40% for the most highly involved items on these population modes, which is in a similar range as reported employing a similar approach in the adult UK Biobank sample (16). For visualization purposes, we produced scatterplots using the highest-loading variables for each mode, color-coded by each individual's score on the respective modes (*SI Appendix*, Fig. S6).

Consistency across Sex and Race/Ethnicity. To assess the degree of similarity of the patterns across the sexes, we split the CCA–ICA subject weights by sex (*SI Appendix*, Figs. S7–S9) and compared sex-specific subject-weight-with-variable correlations to those estimated for the full analysis. Correlations for the three modes ranged between $r = 0.95$ and $r = 1$. We also reestimated the PCA, CCA, and ICA in males and derived out-of-sample canonical correlations for females and vice versa (*SI Appendix*, Figs. S10 and S11), with no sibling/twin pairs included. The aim of this work was not to make comparisons of population subgroups but to detect general population patterns. Since many of the included indicators relating to inequality and socioeconomics are known to differ between ethnic minority groups, we did not regress these variables out of the data. To show that the detected patterns are generalizable we computed subject-weight-with-variable correlations for groups based on parent-ascribed race/ethnicity (*SI Appendix*, Figs. S12–S14) excluding those with a frequency <5% of the total sample (retaining "black," "white," and "other") and compared these to the full analysis. Correlations for the three modes ranged between $r = 0.76$ and $r = 1$. Canonical correlations plotted separately for these groups are shown in *SI Appendix*, Fig. S15. We also performed a 10-fold cross-validation stratified by race, in which no sibling/twin pairs were included, and separately computed average canonical correlations for "black," "white," and "other." These results suggest that the modes capture variance both within and between race/ethnicity (*SI Appendix*, Fig. S16). These results indicate that the patterns are generalizable across sexes and ethnicity/race.

Siblings/Twins. In addition to the abovementioned tests of generalizability across sex, race/ethnicity, and scanner/site, where no siblings/twin pairs were included, confirming that the results are not driven by genetic relatedness, we also plotted the canonical correlations for nonsiblings/siblings/MZ twins (*SI Appendix*, Fig. S17), which confirm that the patterns are present across these subgroups.

Consistency across Sites/Scanners. To further assess the generalizability and robustness of the CCA–ICA patterns, we computed site/scanner-wise CCA–ICA variable correlations and performed correlations comparing these to the full model (*SI Appendix*, Figs. S18–S20). The pattern of the three modes are mostly consistent across sites but with some sites deviating more from the full analysis modes than others ($r = 0.85$ to $r = 0.22$). We also performed a 10-fold cross-validation stratified by scanner (*SI Appendix*, Fig. S21), and leave-one-scanner-out and leave-one-scanner-model-out cross-validations (*SI Appendix*, Figs. S22 and S23), in which no sibling/twin pairs were included, which confirmed that the patterns are generalizable across sites and scanners. All imaging phenotypes were adjusted for site; however, ABCD collects data at 21 sites across the continental United States (https://abcdstudy.org/about), and population-level demographical differences are expected.

Alternative Approach Without Imputation. To ensure that the results were not affected by the imputation procedure for missing data points, we also used an alternative and previously described approach (2) in which we estimated the subject × subject covariance matrix, ignoring missing values, before projecting this approximated covariance matrix to the nearest positive-definite covariance matrix using the MATLAB tool nearestSPD (https://www.mathworks.com/matlabcentral/fileexchange/42885-nearestspd), thereby avoiding the need for imputation of missing values. The correlations between the CCA scores for the first three modes between the original analysis using imputation and this approach were $r = 0.99$, $r = 0.96$, and $r = 0.96$.

Alternative Number of PCA Components. To investigate the impact of choosing a stricter criterion of inclusion of PCA we reran the analysis with 100 PCA components, and compared the resulting CCA scores for the first three modes with those from the original analysis, yielding correlations of $r = 0.98$, $r = 0.91$, and $r = 0.91$.

Adjusting Data for age^2. To address the possibility of nonlinear relationships between age and the various demographic, clinical, and MRI measures features we reran analysis with age^2 added along with the original confound variables and compared the resulting CCA–ICA subject weights for the first three modes with those from the original analysis, yielding correlations of $r = 0.99$, $r = 0.96$, and $r = 0.97$.

Accounting for Motion. To assess the impact of image quality on the results, we used the mean framewise displacement for the DWI as well as for resting state fMRI (rest-fMRI, available for $n = 7{,}493$) images as proxies (27) and correlated these measures with the canonical variates. Correlations with the included behavioral/MRI canonical scores were modest for both DWI motion (variate pair 1, $r = 0.11/r = 0.13$; variate pair 2, $r = -0.02/r = -0.02$; variate pair 3, $r = -0.02/r = -0.03$) and rest-fMRI motion (variate pair 1, $r = 0.16/r = 0.16$; variate pair 2, $r = -0.05/r = -0.05$; variate pair 3, $r = -0.02/r = -0.04$), and the highest motion correlation was for DWI motion with MRI variate 4 ($r = 0.22$) which was excluded after statistical correction (*SI Appendix*, Figs. S24 and S25). Correlations with motion for the CCA–ICA subject weights for DWI/rest-fMRI motion were $r = -0.02/0.04$, $r = 0.13/0.18$, and $r = 0.04/0.09$ for modes 1 to 3, respectively (*SI Appendix*, Tables S6 and S7). Correlations between unadjusted and DWI/rest-fMRI motion-adjusted CCA–ICA subject weights were all $r > 0.98$. Adjusting CCA–ICA subject weights for DWI motion did not change its correlations with the included variables (all $r > 0.99$ between original and adjusted results; *SI Appendix*, Figs. S26–S28).

Data Availability. The data, as well as release notes including documentation of measures, scanning protocols, and imaging QC, are publicly available and can be accessed using the following National Institute of Mental Health data archive: http://dx.doi.org/10.15154/1503209. Code for running the analysis is publicly available at the following Open Science Framework repository: https://doi.org/10.17605/OSF.IO/M39YK.

Results

We identified three distinct modes of covariation, linking brain features to perinatal and early life events, sociocognitive factors, and urbanicity (Fig. 1). Canonical correlations for the first three modes were significant and robust as assessed by 10-fold cross-validation and permutation (out-of-sample $r = 0.61$, $r = 0.42$, and $r = 0.38$, all permuted-$P = 0.001$; *SI Appendix*, Fig. S3).

Mode 1 links perinatal factors and obstetric complications to cognitive ability and brain morphology in late childhood (Fig. 2). Having a twin, premature birth, birth complications requiring

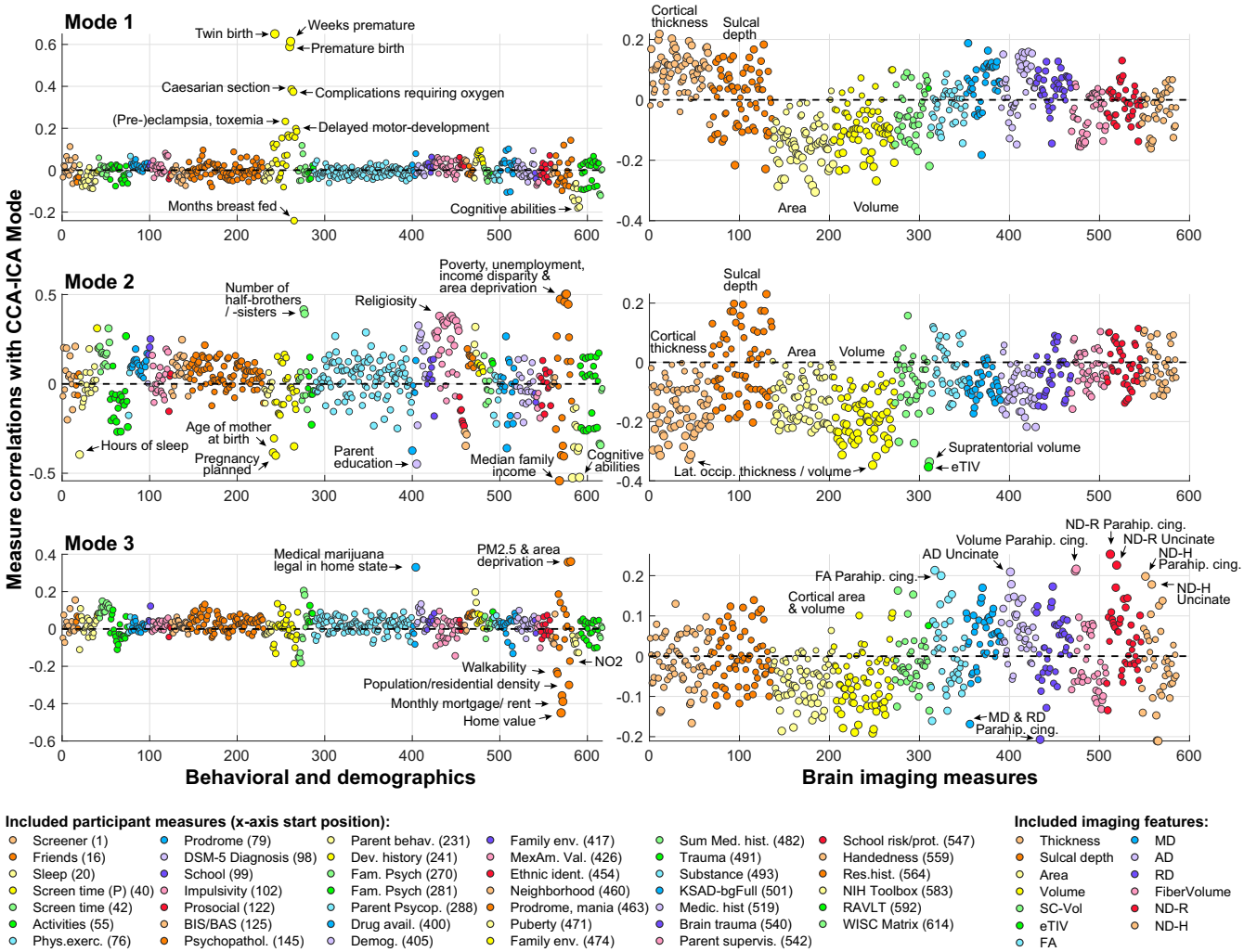

Fig. 1. Rows represent CCA–ICA modes 1–3. Each mode represents an association between a linear combination of behavioral measures (*Left*) with a separate combination of imaging features (*Right*). *x* axes represent the numbered behavioral measures/imaging features. Behavioral measures legend shows starting location on *x* axis for each measure (in parentheses). *y* axes show the correlation between each included variable with the CCA–ICA subject weights. DWI measures abbreviations: fractional anisotropy (FA); mean diffusivity (MD); axial diffusivity (AD); radial diffusivity (RD); neurite density (ND).

oxygen, Caesarian section, (pre)-eclampsia, toxemia, and jaundice are associated with shorter duration of breastfeeding, parent-reported delayed motor development, and lower cognitive scores and linked to a pattern of cortical morphometry and white matter diffusion measures in several brain regions, with lower cortical volume and area and higher thickness in middle temporal, lateral orbitofrontal, and inferior parietal cortex among the highest-loading imaging features.

Mode 2 captures a pattern of economic deprivation and poverty, with the highest loading measures being related to the area deprivation index, such as parent unemployment, neighborhood median household income, income disparity, and violence (Fig. 3). The mode links these measures to lower maternal age at child birth; lower parent education level; unplanned pregnancy; shorter duration of breastfeeding; higher number of half-siblings; higher levels of religiosity; and the child having fewer nightly sleep hours on average, lower grades in school, and worse performance on cognitive tests, jointly forming a dimension of socio-cognitive stratification. This dimension is associated with lower cortical thickness, area, and volume, with total volume, lateral occipital cortical volumes and thickness, and bilateral lingual thickness among the highest-loading imaging features.

Mode 3 links higher air particle matter ($PM_{2.5}$) and area deprivation to lower population density, lower levels of NO_2, lower neighborhood walkability, lower home value and rent but higher home ownership percentage, higher number of half-siblings, and living in a state which has not legalized marijuana for medical use (as of 2016). The mode (Fig. 4 and *SI Appendix*, Table S5) is further associated with reporting emerging signs of puberty such as body hair and with lower area and volumes across the cortex, as well as with white matter indices such as fractional anisotropy, radial diffusivity, neurite density, and tract volumes, with the highest loading measures being related to the parahippocampal cingulum, the uncinate fasciculus, and corpus callosum.

Discussion

Adolescence is a transition period between childhood and adulthood, associated with heightened sensitivity to the social and cultural environment (13). While for most individuals the transition results in successful acquirement of skills and coping strategies required for adulthood and subsequent independence from caregivers, it also coincides with increased risk for mental health issues and psychological madadjustment (14). Research addressing

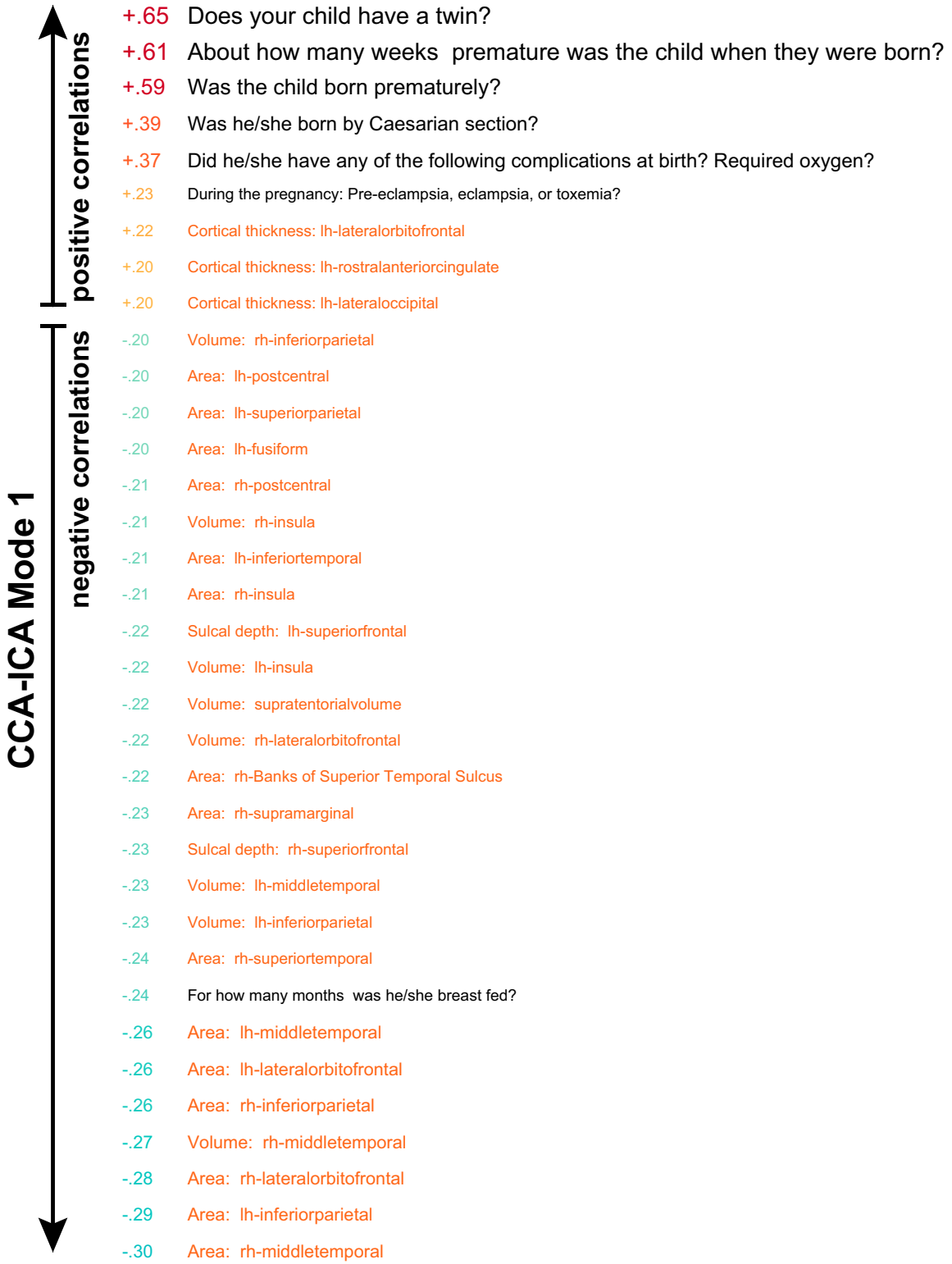

Fig. 2. Mode 1 links obstetric and perinatal complications to cortical area, volume, and thickness. Numbers on the left are correlations between each participant measure and MRI feature with mode 1 CCA–ICA subject weights. Hot and cold colors represent positive and negative correlations, respectively. Text on the right represents the participant measures (black) and imaging features (orange). The top 35 items are shown; for a full list of all measures and their respective correlations, see *SI Appendix*, Table S3.

the social, economic, and environmental conditions affecting adolescent development, facilitating health, and leading to fulfilling adult lives is therefore critical. Here we discuss three modes of population covariation, each linking behavioral, clinical, psychosocial, socioeconomic, and demographical measures to neuroimaging in 7,577 children aged 9 to 11 y.

The first mode links obstetric complications and early life factors such as duration of breastfeeding and motor development,

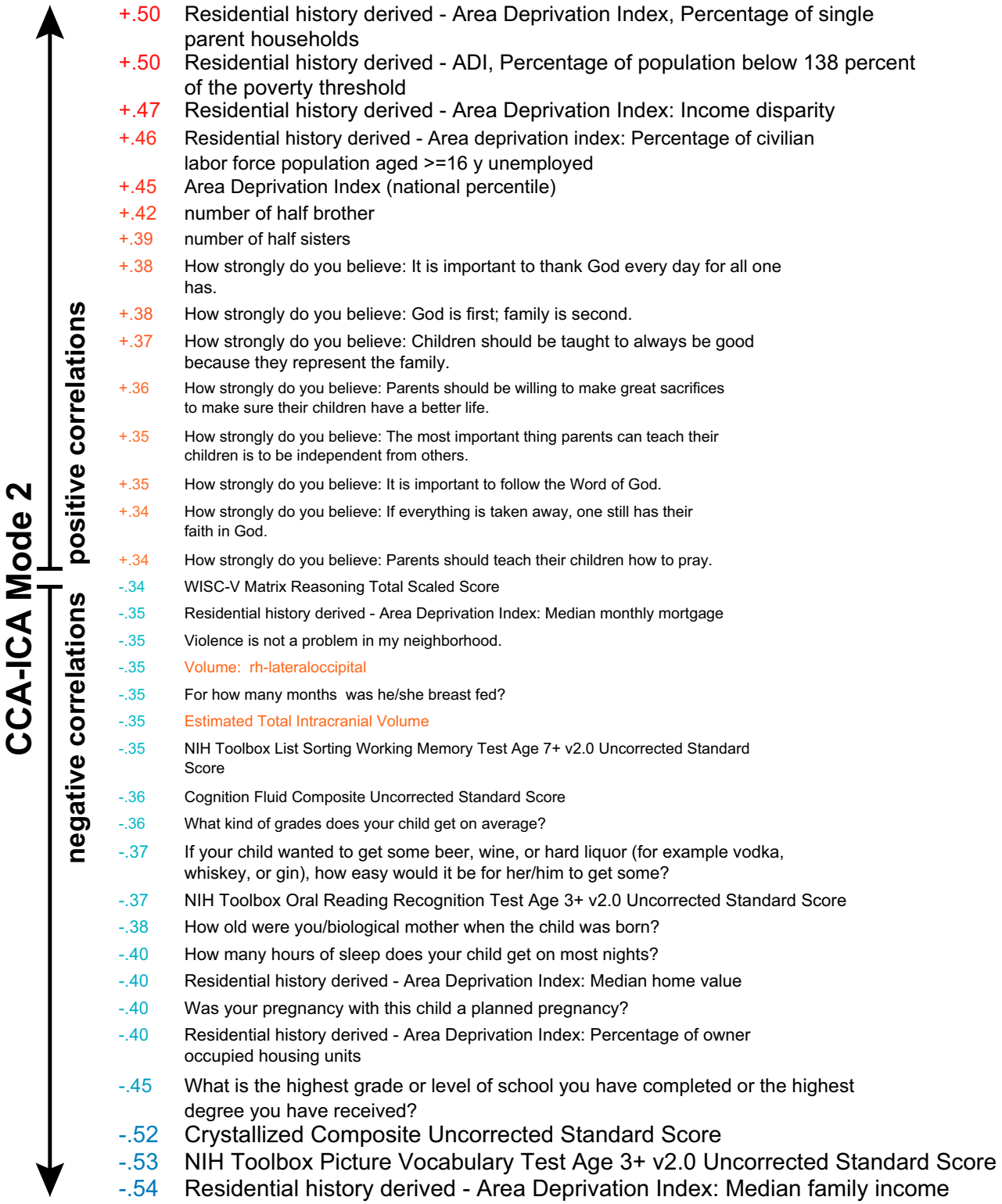

Fig. 3. Variables with the highest correlations with CCA–ICA mode 2 subject weights are shown, forming a dimension of sociocognitive stratification, which is associated with total intracranial volume and regional cortical volumes and thickness. Numbers on the left are correlations; hot and cold colors represent positive and negative correlations, respectively; and arrows indicate increasing positive/negative correlations. Text on the right represents the behavioral measures (black) and imaging features (orange). The top 35 items are shown; for a full list of all measures and associated correlations, see *SI Appendix*, Table S4.

with cognitive ability, cortical surface area, thickness, and volume in late childhood. Preterm birth and obstetric complications are associated with accelerated neonatal brain growth (28), linked to cognitive deficits in childhood and prolonged cognitive maturation into adolescence (29). Deficits may persist through neurodevelopment and into adulthood (30, 31) and constitute risk

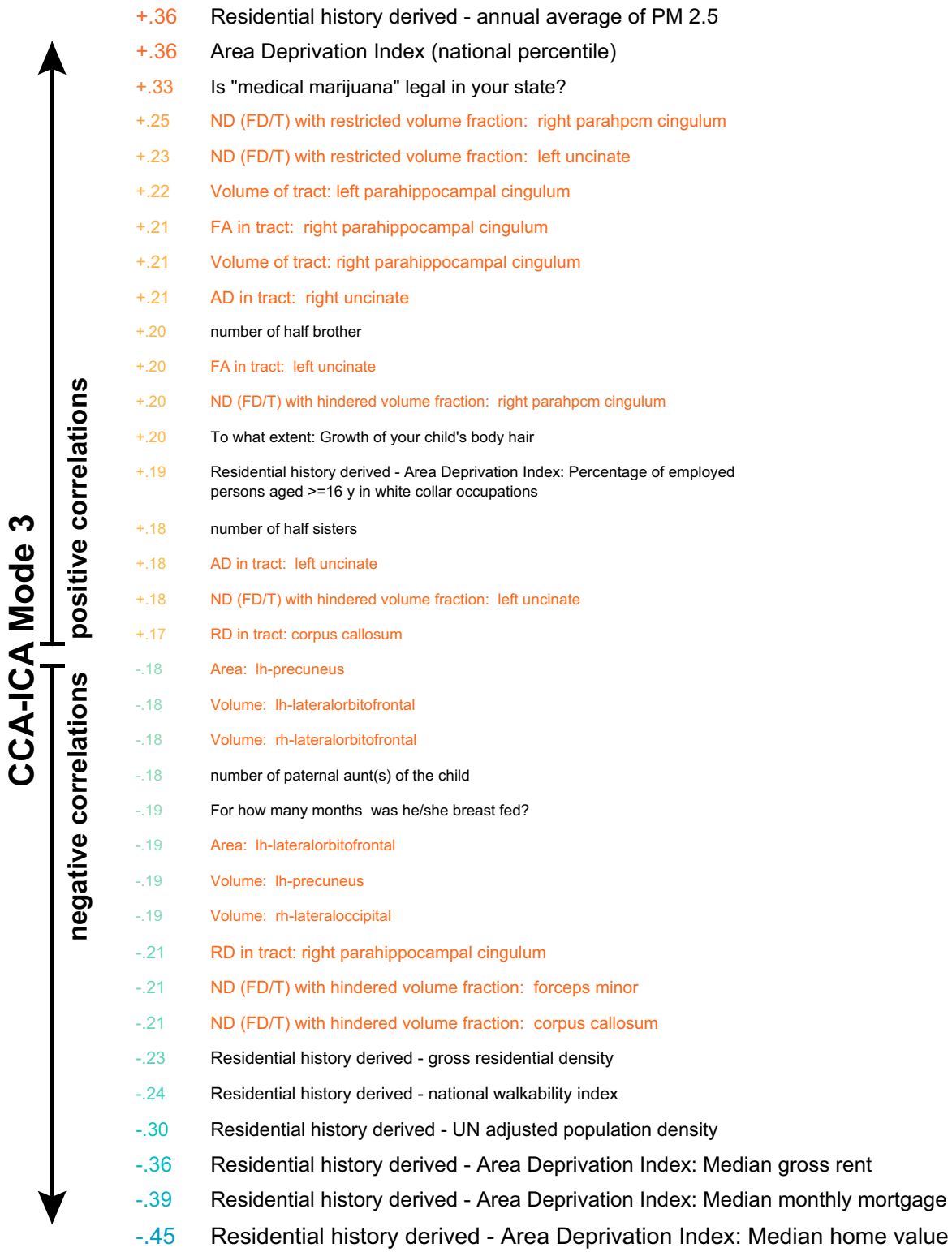

Fig. 4. The variables with the highest correlations with CCA–ICA mode 3 subject weights, linking air pollution, area deprivation, walkability, and population density to brain white matter indices. Numbers on the left are correlations; hot and cold colors represent positive and negative correlations, respectively; and arrows indicate increasing positive/negative correlations. Text on the right represents the behavioral measures (black) and imaging features (orange). The top 35 items are shown; for a full list of all measures and associated correlations, see *SI Appendix*, Table S5.

factors for mental disorders (32, 33). The present results support that children with a history of obstetric and perinatal complications show delayed brain development in late childhood and are consistent with reports associating birth weight with cortical area and brain volume in childhood and adolescence (34), underscoring the importance of taking perinatal factors into account when studying child and adolescent brain development.

The second mode captures a sociocognitive stratification pattern associated with brain volume and regional measures of cortical thickness, area, and volume. Conceptually, the mode shares similarities with a positive–negative population mode linked to brain functional connectivity (2) and structure (17) in adults. The mode links several positive and negative life events and environmental circumstances, with the highest loading factors being related to socioeconomic status, such as poverty, parent unemployment, and education level. It further captures several factors known to be related to social deprivation, such as degree of family planning and early pregnancies, neighborhood level of violence, and level of religious beliefs. These constitute important environmental conditions for neurodevelopment that these children receive from their parents, their community, and society at large. Consistent with the literature on the effect of socioeconomic deprivation on child development, this mode is also associated with less sleep, worse school performance, and lower cognitive ability (35). Cognitive ability is moderately heritable in childhood and adolescence (10, 36, 37), and this likely partly explains the association between child cognition, academic performance, and SES (38). However, the effects of poverty, low socioeconomic status, and early life adversity on brain and cognitive development (20, 39) also underscore the role and importance of social policies aimed at reducing disparities (40) that put some children at a disadvantage, often with lifelong consequences for opportunities, mental and physical health, and quality of life. The neurotypical developmental trajectory at this age is characterized by apparent cortical thinning, likely partly reflecting synaptic pruning (41) and myelination (42). Thicker cortex with higher SES is consistent with reports of accelerated brain maturation in children from low-SES families (43, 44). Indeed, across species and in humans, early life adversity is associated with accelerated maturation of neural systems, possibly at a cost of increased risk for later mental health problems (45).

The third mode reflects an inverse association between particulate matter air pollution ($PM_{2.5}$) and area deprivation on one side and home value, walkability, and population density on the other. Specific geographical information for this mode cannot be discerned, since the ABCD does not provide geographical data about its participants; however, this mode fits a known socioeconomic settlement pattern: low-pollution, high-walkability "sweet spot" neighborhoods in urban areas are typically skewed toward higher-SES households, contrasted with high-pollution, low-walkability "sour spot" neighborhoods associated with lower income (46). Fine particulate matter has harmful effects on brain and cognitive health in both adults (47) and children (48), with white matter injury suggested as a putative mechanism (49). Exposure to $PM_{2.5}$ is associated with adverse health outcomes and disproportionally affects lower-income households (50), underscoring the importance of mapping environmental influences on brain development. Interestingly, this pattern was associated with the legal status of medical marijuana (as of 2016), possibly indicative of geographical differences for this pattern across the US states. Here we document an association with cortical area and volumes, as well as associations with diffusion properties of brain white matter pathways, in particular the parahippocampal cingulum, uncinate fasciculus, corpus callosum, and forceps minor.

While these population level patterns are highly interesting, the cross-sectional and the nonexperimental design warrant caution. People and their brains, genes, and environments are not varying randomly but are highly correlated (51), likely along multiple dimensions, which complicates causal and mechanistic inference. This is especially relevant for population-based neuroimaging, in which subtle confounds can induce spurious associations (5). These general caveats notwithstanding, these valuable resources represent an unprecedented opportunity to reveal covarying patterns of sociodemographics, cognitive abilities, mental health, and brain imaging data, beyond simple bivariate associations, which are potentially highly informative of the biology, psychology, and sociology of childhood and adolescent brain development and psychological adaptation.

In contrast to the standard regression approach which models one outcome variable at a time and typically includes only a few covariates, the combined multivariate approach employed here considers the full pattern of covariability between variables. Our approach is therefore well suited for capturing population patterns by maximizing statistical power. However, it does not allow for interpretation of specific associations between pairs of variables. Overfitting can be a challenge with multivariate approaches, in particular in small samples and for complex models (52). Currently, there are no comparable samples to ABCD in which to independently assess the generalizability of the results. However, the current results were obtained in a large sample, using data reduction as well as 10-fold cross-validation with all relevant analysis steps performed within the cross-validation and permutation loop, to avoid overfitting and assess generalizability. All of the patterns are purely correlational and also treated analytically and reported as such. It is also entirely possible, and highly probable, that these patterns are further correlated with other important phenomena not measured or included in the current analysis. In-scanner subject motion affects image quality and can influence associations between imaging and nonimaging variables, and taking this into account is particularly important in child and adolescent samples (27). While we cannot completely rule out motion as a confound, our follow-up analysis suggests that the detected patterns cannot simply be attributed to individual differences in subject motion. The current approach also effectively captures differential patterns involving the same measures. For example, higher cortical thickness, indicative of delayed maturation, is independently associated with sociocognitive stratification, higher cognitive ability, and SES, as well as with obstetric and perinatal complications, lower cognitive ability, and delayed speech and motor development. Another example is duration of breastfeeding, which was independently associated both with obstetric complications and with sociocognitive stratification, associated with differential patterns of brain differences. It is also worth noting that for all of the three brain–behavior patterns, variables related to the environments the children inhabit, at different temporal and geographical scales, are among the highest loading features. These children will be followed up with new brain scans every second year for the next 10 y, and assessing the predictive value of these patterns for later maladjustment and mental illness as the participant transition from childhood into adolescence and early adulthood will be important follow-ups to this study. These independent and coexisting associations with brain structure emphasize the importance of multidimensional considerations for understanding child and adolescent neurodevelopment and support that political priorities and decisions aiming to improve health outcomes and adaptation during transformative life phases should be based on interdisciplinary perspectives integrating social, psychological, and biological sciences (53).

ACKNOWLEDGMENTS. D.A. is funded by the South-Eastern Norway Regional Health Authority (grants 2019107 and 2020086). T.K. is funded by the Research Council of Norway (grant 276082). A.F.M. gratefully acknowledges support from the Dutch Organisation for Scientific Research via a Vernieuwingsimpuls VIDI fellowship (016.156.415) and a Wellcome Trust Innovator award (215698/Z/19/Z). S.M.S. is funded by Wellcome Trust grants (203139/Z/16/Z, 098369/Z/12/Z, and 215573/Z/19/Z). L.T.W. is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research

and innovation program (ERC Starting Grant 802998); the Research Council of Norway (grants 249795, 298646, and 300768); the South-East Norway Regional Health Authority (grant 2019101); and the Department of Psychology, University of Oslo.

1. H. Tost, F. A. Champagne, A. Meyer-Lindenberg, Environmental influence in the brain, human welfare and mental health. *Nat. Neurosci.* **18**, 1421–1431 (2015).
2. S. M. Smith *et al*., A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat. Neurosci.* **18**, 1565–1567 (2015).
3. B. J. Casey *et al*.; ABCD Imaging Acquisition Workgroup, The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).
4. M. P. Paulus, W. K. Thompson, The challenges and opportunities of small effects: The new normal in academic psychiatry. *JAMA Psychiatry* **76**, 353–354 (2019).
5. S. M. Smith, T. E. Nichols, Statistical challenges in "big data" human neuroimaging. *Neuron* **97**, 263–268 (2018).
6. E. S. Finn *et al*., Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* **18**, 1664–1671 (2015).
7. T. Kaufmann *et al*., Delayed stabilization and individualization in connectome development are related to psychiatric disorders. *Nat. Neurosci.* **20**, 513–515 (2017).
8. T. Kaufmann *et al*.; Karolinska Schizophrenia Project (KaSP), Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nat. Neurosci.* **22**, 1617–1623 (2019).
9. K. Franke, C. Gaser, Ten years of *BrainAGE* as a neuroimaging biomarker of brain aging: What insights have we gained? *Front. Neurol.* **10**, 789 (2019).
10. D. Alnæs *et al*., Association of heritable cognitive ability and psychopathology with white matter properties in children and adolescents. *JAMA Psychiatry* **75**, 287–295 (2018).
11. K. Watanabe *et al*., A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**, 1339–1348 (2019).
12. P. M. Visscher *et al*., 10 years of GWAS discovery: Biology, function, and translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
13. S.-J. Blakemore, K. L. Mills, Is adolescence a sensitive period for sociocultural processing? *Annu. Rev. Psychol.* **65**, 187–207 (2014).
14. T. Paus, M. Keshavan, J. N. Giedd, Why do many psychiatric disorders emerge during adolescence? *Nat. Rev. Neurosci.* **9**, 947–957 (2008).
15. A. Caspi *et al*., The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clin. Psychol. Sci.* **2**, 119–137 (2014).
16. K. L. Miller *et al*., Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).
17. A. Llera, T. Wolfers, P. Mulders, C. F. Beckmann, Inter-individual differences in human brain structure and morphology link to variation in demographics and behavior. *eLife* **8**, e44443 (2019).
18. J. Bor, G. H. Cohen, S. Galea, Population health in an era of rising income inequality: USA, 1980-2015. *Lancet* **389**, 1475–1490 (2017).
19. N. L. Hair, J. L. Hanson, B. L. Wolfe, S. D. Pollak, Association of child poverty, brain development, and academic achievement. *JAMA Pediatr.* **169**, 822–829 (2015).
20. K. G. Noble *et al*., Family income, parental education and brain structure in children and adolescents. *Nat. Neurosci.* **18**, 773–778 (2015).
21. C. Leys, C. Ley, O. Klein, P. Bernard, L. Licata, Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **49**, 764–766 (2013).
22. D. J. Hagler Jr. *et al*., Image processing and analysis methods for the Adolescent Brain Cognitive Development study. *Neuroimage* **202**, 116091 (2019).
23. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
24. W. J. Krzanowski, *Principles of Multivariate Analysis: A User's Perspective*, W. J. Krzanowski, Ed. (Oxford University Press, Inc., 1988).
25. A. M. Winkler, G. R. Ridgway, M. A. Webster, S. M. Smith, T. E. Nichols, Permutation inference for the general linear model. *Neuroimage* **92**, 381–397 (2014).
26. A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* **10**, 626–634 (1999).
27. A. Alexander-Bloch *et al*., Subtle in-scanner motion biases automated measurement of brain anatomy from in vivo MRI. *Hum. Brain Mapp.* **37**, 2385–2397 (2016).
28. R. C. Knickmeyer *et al*., Impact of demographic and obstetric factors on infant brain volumes: A population neuroscience study. *Cereb. Cortex* **27**, 5616–5625 (2017).
29. R. Everts, C. G. Schöne, I. Mürner-Lavanchy, M. Steinlin, Development of executive functions from childhood to adolescence in very preterm-born individuals—Alongitudinal study. *Early Hum. Dev.* **129**, 45–51 (2019).
30. L. D. Breeman, J. Jaekel, N. Baumann, P. Bartmann, D. Wolke, Preterm cognitive function into adulthood. *Pediatrics* **136**, 415–423 (2015).
31. E. A. Figueiró-Filho *et al*., Neurological function in children born to preeclamptic and hypertensive mothers—A systematic review. *Pregnancy Hypertens.* **10**, 1–6 (2017).
32. G. Mezquida *et al*., Obstetric phenotypes in the heterogeneity of schizophrenia. *J. Nerv. Ment. Dis.* **206**, 882–886 (2018).
33. R. M. Murray, V. Bhavsar, G. Tripoli, O. Howes, 30 years on: How the neurodevelopmental hypothesis of schizophrenia morphed into the developmental risk factor model of psychosis. *Schizophr. Bull.* **43**, 1190–1196 (2017).
34. K. B. Walhovd *et al*.; Pediatric Imaging, Neurocognition, and Genetics Study, Long-term influence of normal variation in neonatal characteristics on human brain development. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 20089–20094 (2012).
35. L. E. Philbrook, J. B. Hinnant, L. Elmore-Staton, J. A. Buckhalt, M. El-Sheikh, Sleep and cognitive functioning in childhood: Ethnicity, socioeconomic status, and sex as moderators. *Dev. Psychol.* **53**, 1276–1285 (2017).
36. A. G. Allegrini *et al*., Genomic prediction of cognitive traits in childhood and adolescence. *Mol. Psychiatry* **24**, 819–827 (2019).
37. R. Plomin, I. J. Deary, Genetics and intelligence differences: Five special findings. *Mol. Psychiatry* **20**, 98–108 (2015).
38. M. Trzaskowski *et al*., Genetic influence on family socioeconomic status and children's intelligence. *Intelligence* **42**, 83–88 (2014).
39. S. B. Johnson, J. L. Riis, K. G. Noble, State of the art review: Poverty and the developing brain. *Pediatrics* **137**, e20153075 (2016).
40. A. S. Hermansen, N. T. Borgen, A. Mastekaasa, Long-term trends in adult socioeconomic resemblance between former schoolmates and neighbouring children. *Eur. Sociol. Rev.*, 10.1093/esr/jcz066 (2019).
41. J. H. Gilmore, R. C. Knickmeyer, W. Gao, Imaging structural and functional brain development in early childhood. *Nat. Rev. Neurosci.* **19**, 123–137 (2018).
42. V. S. Natu *et al*., Apparent thinning of human visual cortex during childhood is associated with myelination. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20750–20759 (2019).
43. K. Z. LeWinn, M. A. Sheridan, K. M. Keyes, A. Hamilton, K. A. McLaughlin, Sample composition alters associations between age and brain structure. *Nat. Commun.* **8**, 874 (2017).
44. L. R. Piccolo, E. C. Merz, X. He, E. R. Sowell, K. G. Noble; Pediatric Imaging, Neurocognition, Genetics Study, Age-related differences in cortical thickness vary by socioeconomic status. *PLoS One* **11**, e0162511 (2016).
45. B. L. Callaghan, N. Tottenham, The stress acceleration hypothesis: Effects of early-life adversity on emotion circuits and behavior. *Curr. Opin. Behav. Sci.* **7**, 76–81 (2016).
46. J. D. Marshall, M. Brauer, L. D. Frank, Healthy neighborhoods: Walkability and air pollution. *Environ. Health Perspect.* **117**, 1752–1759 (2009).
47. E. H. Wilker *et al*., Long-term exposure to fine particulate matter, residential proximity to major roads and measures of brain structure. *Stroke* **46**, 1161–1166 (2015).
48. L. Calderón-Garcidueñas *et al*., Exposure to severe urban air pollution influences cognitive outcomes, brain volume and systemic inflammation in clinically healthy children. *Brain Cogn.* **77**, 345–355 (2011).
49. R. M. Babadjouni *et al*., Clinical effects of air pollution on the central nervous system; a review. *J. Clin. Neurosci.* **43**, 16–24 (2017).
50. B. Bowe, Y. Xie, Y. Yan, Z. Al-Aly, Burden of cause-specific mortality associated with PM2.5 air pollution in the United States. *JAMA Netw. Open* **2**, e1915834 (2019).
51. E. Turkheimer, "Genome wide association studies of behavior are social science" in *Philosophy of Behavioral Biology*, K. S. Plaisance, T. A. C. Reydon, Eds. (Springer Science, New York, 2012), pp. 43–64.
52. R. A. Poldrack, G. Huckins, G. Varoquaux, Establishment of best practices for evidence for prediction: A review. *JAMA Psychiatry*, 10.1001/jamapsychiatry.2019.3671 (2019).
53. M. J. Farah, Socioeconomic status and the brain: Prospects for neuroscience-informed policy. *Nat. Rev. Neurosci.* **19**, 428–438 (2018).