Biological Psychiatry: CNNI

Archival Report

Manifold Learning Uncovers Nonlinear Interactions Between the Adolescent Brain and Environment That Predict Emotional and Behavioral Problems

Erica L. Busch, May I. Conley, and Arielle Baskin-Sommers

ABSTRACT

BACKGROUND: To progress adolescent mental health research beyond our present achievements—a complex account of brain and environmental risk factors without understanding neurobiological embedding in the environment—we need methods to uncover relationships between the developing brain and real-world environmental experiences.

METHODS: We investigated associations between brain function, environments, and emotional and behavioral problems using participants from the Adolescent Brain Cognitive Development (ABCD) Study (n = 2401 female). We applied manifold learning, a promising technique for uncovering latent structure from high-dimensional biomedical data such as functional magnetic resonance imaging. Specifically, we developed exogenous PHATE (potential of heat-diffusion for affinity-based trajectory embedding) (E-PHATE) to model brain-environment interactions. We used E-PHATE embeddings of participants' brain activation during emotional and cognitive processing tasks to predict individual differences in cognition and emotional and behavioral problems both cross-sectionally and longitudinally.

RESULTS: E-PHATE embeddings of participants' brain activation and environments at baseline showed moderate-to-large associations with total, externalizing, and internalizing problems at baseline, across several subcortical regions and large-scale cortical networks, compared with the zero-to-small effects achieved by voxelwise data or common low-dimensional embedding methods. E-PHATE embeddings of the brain and environment at baseline were also related to emotional and behavioral problems 2 years later. These longitudinal predictions showed a consistent moderate effect in the frontoparietal and attention networks.

CONCLUSIONS: The embedding of the adolescent brain in the environment yields enriched insight into emotional and behavioral problems. Using E-PHATE, we demonstrated how the harmonization of cutting-edge computational methods with longstanding developmental theories advances the detection and prediction of adolescent emotional and behavioral problems.

https://doi.org/10.1016/j.bpsc.2024.07.001

Nearly 75% of all mental health disorders onset during adolescence, with half of all mental health disorders occurring by age 14 (1). Adolescents who experience mental health problems are at heightened risk for lifelong challenges including lower educational attainment, increased involvement with the legal system, and chronic physical and mental health problems (2). Given the impacts of mental health problems on individuals and society, developmental scientists have long grappled with understanding the emergence of emotional and behavioral problems in youths.

An extensive body of research has identified factors related to emotional and behavioral problems, including neurobiological and environmental factors (3,4). Much of this work has been siloed into work specifying the neurobiology or the environmental factors related to mental health problems in adolescence. Neurobiological theories of emotional and behavior problems have emphasized that 3 key brain regions are especially sensitive during adolescent development: prefrontal cortex, amygdala, and hippocampus (5). These brain regions support self-regulation and affective processing (6,7), and differences in their functional activation have been related to various aspects of emotional and behavioral problems (5). Other research has identified environmental exposures that increase risk for the development of emotional and behavioral problems (8). Meta-analyses report medium to large effects between adversity in adolescents' families (e.g., conflict, caregiver nonacceptance) and neighborhoods (e.g., experiencing violence or disadvantage) and emotional and behavioral problems (9–11).

Some environmental risk factors (e.g., parenting styles, community disadvantage) have been related to the function of mental health–related brain regions (12,13). For example, a

ISSN: 2451-9022

recent study found that the interaction of neighborhood adversity and lower executive network activation during an emotional working memory task was related to higher externalizing problems in adolescents (14). Additional work found that the interaction between neighborhood adversity and decreased amygdala activation during an emotional introspection task was related to higher externalizing problems in a sample of adolescents of Mexican origin (15). Furthermore, another study found that neighborhood and family adversity interacted with prefrontal cortex functional connectivity to predict internalizing symptoms (16). Across these studies, we can stitch together a model of emotional and behavioral problems that includes interactions between experiences in adolescents' environments and brain function in regions involved in emotion processing and cognition. Most previous research has modeled this interaction as a linear combination between univariate measures of brain and environment—in other words, considering a single measurement of environment and a univariate signal of brain activation. Some recent work has used multivariate approaches, such as canonical correlation analysis, partial least squares regression, or principal component analysis (PCA) ridge regression, to maximize brain-behavior associations (17–21). However, these approaches focus on learning components of a multivariate neural representation that are maximally predictive of a target behavior, which may not necessarily reflect the intrinsic geometry of brain activation itself that could arise through unsupervised methods. Moreover, they frequently fail to account for the multivariate measures of the environment in conjunction with brain data.

Studying the nonlinear, multidimensional interplay between the adolescent brain and environmental risks requires computational methods to combine and reveal structure in high-dimensional, multimodal data. Manifold learning is increasingly popular for highlighting complex latent structure in high-dimensional biological data (22). Manifold learning is an unsupervised, data-driven approach in which the dimensionality reduction step is not optimized to maximize a prediction, but rather discovers a manifold of given data. Downstream, manifold embeddings can be used to test associations between the manifold and other information and gauge the quality or type of information represented within the manifold. The algorithm PHATE (potential of heat-diffusion for affinity-based trajectory embedding) was specifically designed for high-dimensional, noisy biomedical data and has been applied in prior work to uncover local and global latent structure in functional magnetic resonance imaging (fMRI) data (23–25). Prior research showed that combining PHATE with additional data (e.g., temporal dynamics of brain responses) enhances the relevance of embeddings for understanding complex cognitive processing (i.e., during movie viewing) (26). However, standard PHATE implementations cannot account for interactions of additional variables outside of the high-dimensional input data (e.g., brain data).

Here, using the Adolescent Brain Cognitive Development Study (ABCD Study) baseline sample and 2-year follow-up data, we investigated the interplay of environment and brain function on emotional and behavioral problems. We tested 1) whether PHATE can be used to enhance the behavioral relevance of task-based, developmental fMRI data and 2) whether an updated version of PHATE can be combined with environmental data to discover latent geometric structure connecting adolescents' brains, environments, and emotional and behavioral problems. First, as a proof of concept, we showed that PHATE embeddings of brain activation during cognitive and emotion processing (27) were strongly associated with individual differences in working memory performance in 9- to 10-year-old adolescents. Next, we combined the PHATE brain activation manifold with measurements of adolescents' environments into a multiview manifold. Multiview approaches combine different measurements collected from the same samples into a single representation to be embedded in lower dimensions. For example, temporal PHATE is a recently introduced multiview algorithm that combines 2 signals endogenous to brain data (i.e., calculated directly from the fMRI measurements) (26). In the current study, we introduced exogenous PHATE (E-PHATE), which combines participants' PHATE brain activation manifold with data about the same participants collected externally (i.e., family and neighborhood adversity). We hypothesized that the development of emotional and behavioral problems would be related to a nonlinear interaction between the adolescent brain and their environments. E-PHATE embeddings showed a stronger relationship with emotional and behavioral problems both cross-sectionally and longitudinally than either the PHATE or the original voxelwise data (3,4,28,29). E-PHATE sheds light on the neural-environment interplay and improves the detection and prediction of emotional and behavioral problems in adolescents.

METHODS AND MATERIALS

Participants

Participants were adolescents included in the ABCD Annual Data Release 4.0 (https://nda.nih.gov/study.html?id=1299). Neuroimaging and environment data were from the baseline assessment (ages 9–10). Emotional and behavioral data were from the baseline and the 2-year follow-up (ages 11–12). Participants were excluded for missing fMRI, environmental, or mental health measures, which resulted in 4732 participants being included in baseline analyses and 2371 participants being included in longitudinal analyses (see Table 1 for sample demographics; Supplemental Methods S1; Tables S3 and S4; and Figure S1 in Supplement 1 for participant selection).

Environment Measures

Measures of the environment were selected to characterize participants' family and neighborhood environments at the baseline time point. Family environment was measured using participants' perception of family threat (anger and conflict expressed among family members) and family support (caregiver acceptance). Neighborhood environment was measured using participant and caregiver assessment of perceived safety/crime, and the area deprivation index, a composite index of neighborhood socioeconomic disadvantage (Supplemental Methods S2 in Supplement 1).

Emotional and Behavioral Measures

Measures of emotional and behavioral problems were assessed using *t* scores from the baseline and 2-year follow-up

Table 1. Demographics for Participants Included in the Baseline (Ages 9–10) and 2-Year Follow-Up (Ages 11–12) Analyses

Time Point	Baseline, n = 4372	2-year Follow-Up, n = 2371
Sex, Female	2401 (50.7%)	1182 (49.9%)
Race/Ethnicity		
Asian	107 (2.26%)	48 (2.0%)
Black	518 (10.9%)	198 (8.4%)
Hispanic or Latinx	883 (18.7%)	424 (17.9%)
Other	513 (10.8%)	243 (10.2%)
White	2710 (57.3%)	1458 (61.5%)
Caregiver Education		
Less than HS diploma	207 (4.4%)	86 (3.6%)
HS diploma/GED	426 (9.0%)	194 (8.2%)
Some college	1299 (27.5%)	653 (27.5%)
Bachelor's degree	1463 (30.9%)	772 (32.6%)
Postgraduate degree	1334 (28.2%)	664 (28.0%)
Unknown	3 (0.06%)	2 (0.08%)
Family Income		
<$50,000	1047 (22.1%)	491 (20.7%)
$50,000–$99,999	1298 (27.4%)	714 (30.1%)
>$100,000	2053 (43.4%)	1017 (42.9%)
Unknown	334 (7.1%)	149 (6.3%)

Values are presented as n (%).
HS, high school; GED, General Educational Development.

data from the Achenbach System of Empirically Based Assessment Child Behavior Checklist (CBCL), which is a 119-item parent-/caregiver-report survey of adolescent emotional and behavioral problems validated for use in children ages 6 to 18 years (30). Primary analyses examined total problems and externalizing and internalizing broadband scales. In supplemental analyses, we examined anxious/depression, withdrawn/depression, somatization, aggression, and rule-breaking behavior syndrome scales.

Neuroimaging Task Data

The in-scanner emotional n-back (EN-back) task was designed to engage emotion and memory processing (27,31). During each fMRI run, participants performed four 0-back (low memory load) and four 2-back (high memory load) blocks with happy, fearful, or neutral face or place stimuli. EN-back performance was measured with sensitivity, calculated as $d' = z(\text{hits}) - z(\text{false alarms})$ and adjusted for extreme values using the Hautus adjustment method (32,33). Task information is detailed further in Supplemental Methods S3 in Supplement 1.

fMRI data were preprocessed by the ABCD Study Data Analysis, Informatics, and Resource Center (34). EN-back fMRI activation was estimated for each participant using general linear models. Following prior studies, cognitive processing activation was measured as the contrast of 2-back and 0-back blocks, and emotion processing activation was measured as the contrast between emotional and neutral face blocks (14,35). Further information about acquisition and processing is provided in Supplemental Methods S4 in Supplement 1.

For each contrast, we analyzed beta weights from cortical networks [defined using the Yeo 7 network solution (36) and the Shaefer 400 parcellation (37)] and subcortical regions [defined using the Scale I Tian subcortical parcellation (38)]. Then, we extracted 1 beta weight per voxel within each region or network and then vectorized the multivoxel beta weights, which resulted in 1 vector of voxelwise beta weights for each region, participant, and contrast. All participants' vectors were stacked into a single matrix for each region and contrast, and henceforth we refer to these matrices as voxelwise data.

PHATE Manifold Learning

We tested whether manifold learning uncovered behaviorally relevant brain activation during cognitive and emotional processing, which could be used to improve prediction of emotional and behavioral problems. First, we applied the PHATE algorithm. PHATE embeddings denoise and highlight local and global nonlinear structure among data points in a low-dimensional representation. Prior work has shown that PHATE embeddings of fMRI data improve the sensitivity of the data for predicting features such as functional brain maturity (25) and visual category information (24,26). Using PHATE, we embedded the voxelwise data for each region and contrast into lower dimensions to test whether the PHATE embeddings uncovered individual differences in brain function related to cognition more strongly than the voxelwise data.

E-PHATE Manifold Learning

We designed E-PHATE to model this interplay as a low-dimensional manifold. We applied a dual-diffusion (26) approach to combine exogenous information about adolescents' environments with their brain activation manifold. The E-PHATE procedure starts by calculating a PHATE diffusion matrix over the voxelwise data, which can be considered the affinity among participants' voxelwise activation vectors. E-PHATE also calculates a second affinity matrix over those participants' scores on additional, exogenous variables (i.e., family conflict, caregiver acceptance, participant- and caregiver-perceived neighborhood crime/safety, and neighborhood disadvantage), and these views are combined using dual-diffusion to calculate the E-PHATE diffusion matrix. This matrix is then embedded into D dimensions using multidimensional scaling (where $D \in \mathbb{N}$; $D \in \{2, 3\}$ for visualization) (Figure 1; see Supplemental Methods S5 in Supplement 1 for the algorithm).

Benchmarking Manifold Learning Methods

In the main analyses, we benchmarked the relevance of E-PHATE and PHATE embeddings against their corresponding (i.e., within region and contrast) voxelwise data. In supplemental analyses, we benchmarked E-PHATE with PCA (a linear dimensionality reduction method) and universal manifold approximation and projection (UMAP) (39) (a common nonlinear embedding method) (Supplemental Methods S7 in Supplement 1). We also benchmarked E-PHATE with variants to test the impact of specific environment variables and algorithmic choices (Supplemental Methods S8 and S9 in Supplement 1). As in prior work, all analyzed embeddings were 20-dimensional for consistency across methods and regions

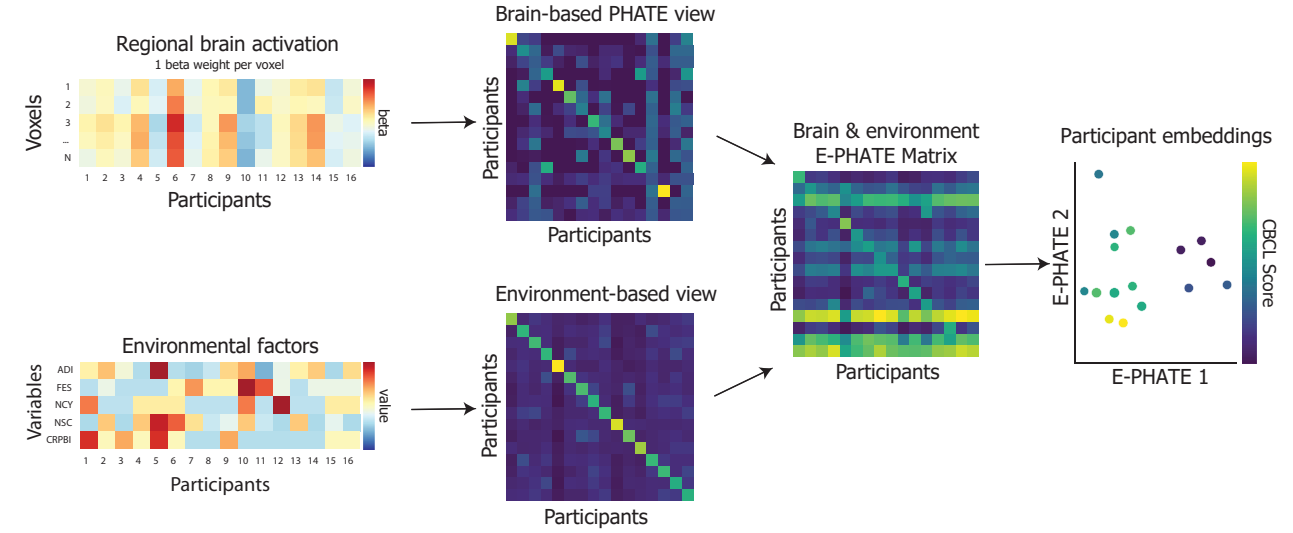

Figure 1. Exogenous PHATE (E-PHATE) procedure. **(A)** E-PHATE models the interactions between brain activation and exogenous information about participants using multiview manifold learning. In this schematic, the first view of E-PHATE takes as inputs a vector of voxelwise beta values for each participant and computes a PHATE-based affinity matrix between participants' brain activations. The second view takes a vector of environment scores for each of those participants and builds an affinity matrix across those scores. Both matrices are row-normalized to become transition probability matrices. These 2 views are combined into the E-PHATE diffusion matrix, which now captures both brain and environmental relations among participants. The E-PHATE matrix is then embedded using metric multidimensional scaling. Two dimensions and a subset of participants are shown for visualization; 20 dimensions were used for the main analysis. Participants' coordinates in E-PHATE dimensions visually reflect individual differences along emotional and behavioral problems (e.g., externalizing problem scores). The ADI is a composite index of neighborhood disadvantage, the CRPBI measures caregiver acceptance, the FES measures family conflict, the NCY measures youth-perceived neighborhood safety/crime, and the NSC measures caregiver-perceived neighborhood safety/crime. **(B)** Analyses are presented using beta values extracted for voxels in the bilateral amygdala and hippocampus and surface vertices in 3 cortical networks: the frontoparietal, dorsal, and ventral attention networks. ADI, area deprivation index; CBCL, Child Behavior Checklist; CRPBI, Child's Report of Parent Behavior Inventory; FES, Family Environment Scale; NCY, Neighborhood Crime, Youth report; NSC, Neighborhood Safety Protocol, Caregiver report.

(23,26). Voxelwise data dimensionality is provided in Table S2 in Supplement 1.

Prediction of Emotional and Behavior Problems and Task Performance From Brain Data

Cross-sectional analyses used participants' brain data at baseline to predict behavioral scores at baseline (n = 4732). Longitudinal analyses used brain data at baseline to predict scores at the 2-year follow-up, controlling for the participant's score at baseline (n = 2371). For each cross-validation fold, multiple linear regressions were trained on 95% of participants' brain data (either voxelwise or embedding) to predict participants' EN-back task or CBCL scores (20,21,26). Regressions were then applied to predict scores from the brain data (i.e., voxelwise or embedding) of held-out participants and scored as the partial Spearman's correlation (ρ) between predicted and true scores for each fold of held-out participants. ρ values were then averaged across folds. In our analyses, we compared the partial Spearman's ρ of multiple linear regression models trained on voxelwise data, PHATE, and E-PHATE embeddings. Covariates included scanner serial number for cross-sectional and longitudinal analyses and participants' baseline scores for the behavioral measure being predicted in longitudinal analyses. The performance of regression models trained on different data representations (e.g., voxelwise vs.

E-PHATE) was compared across representations using pairwise permutation tests (10,000 iterations) and Bonferroni correction; comparisons across representations refererence these p values. Further details about regression models and statistical testing are included in Supplemental Methods S6 and S10 in Supplement 1, respectively. Analysis pipeline is available at: https://github.com/ericabusch/manifold_abcd_psychopathology_bpcnni. Stand-alone software package for E-PHATE is available at: https://github.com/ericabusch/EPHATE.

RESULTS

PHATE Strengthened Associations Between Brain Activation and EN-Back Performance

As a proof of concept to confirm whether manifold learning could improve representation of cognitively relevant brain activity, we first tested the association between standard PHATE embeddings and participants' EN-back task scores. Voxelwise 2-back versus 0-back activation showed moderate associations with EN-back performance ($\rho < 0.20$) in all regions (Table 2). PHATE embeddings of 2-back versus 0-back activation were significantly related to EN-back task performance in all regions (Figure 2 and Table 2; Supplemental Data 2 in Supplement 3) with a large effect size ($\rho > 0.20$). The greatest effects were observed in the frontoparietal and attention networks ($\rho > 0.52$), which were more than double the effect sizes (ρ) for regression models trained on voxelwise data. Given previous research linking frontoparietal and attention networks with higher-order cognitive abilities and working memory (35,40,41), these results demonstrate that PHATE optimized the sensitivity of the fMRI data for detecting brain activation related to cognitive performance.

Consistent with prior research (35), none of the voxelwise data for the emotion versus neutral face contrast significantly related to EN-back performance. In contrast, PHATE embeddings of emotion versus neutral activation showed moderate associations with EN-back performance ($\rho > 0.14$ in frontoparietal and attention networks). Therefore, the PHATE embedding enhanced access to brain activation related to emotion processing during the EN-back task.

E-PHATE Strengthened Cross-Sectional Associations With Emotional and Behavioral Problems

We tested whether E-PHATE strengthened associations between brain function and CBCL scores (30) by comparing the partial Spearman's ρs of multiple linear regression models trained on voxelwise data, PHATE, and E-PHATE embeddings. Voxelwise data from the 2-back versus 0-back working memory contrast showed a small effect that was nonetheless statistically significant: a relationship to CBCL total problems only in the hippocampus (ρ = 0.032, 95% CI, 0.005–0.063) (Table 3). PHATE embeddings of the 2-back versus the 0-back contrast were significantly related to total problems in all regions of interest (ROIs) with small-to-moderate effect sizes, but only outperformed voxelwise data significantly in the frontoparietal network. E-PHATE reflected stronger associations between brain activation and total problems compared with both the voxelwise data and PHATE embeddings for the 2-back versus the 0-back contrast for every region (all corrected ps $< .01$, except frontoparietal E-PHATE vs. PHATE $p < .05$). The magnitude of the associations between E-PHATE embeddings and overall emotional and behavioral problems were similarly moderate-to-high (ρ = 0.143–0.152) across all regions (Figure 3A and Table 3; Supplemental Data 2 in Supplement 3).

Replicating previous research that showed null associations between emotion processing activation and emotional and behavioral problems (42), voxelwise activity from the emotion versus neutral contrast was not significantly related to individual differences in CBCL total problems in any ROI. PHATE embeddings of the emotion versus neutral contrast performed similarly to the voxel data in terms of the strength of its relationship with total problems (all ρs < 0.035 for voxel and PHATE; small effect size). E-PHATE significantly strengthened this relationship compared with the voxelwise data and PHATE embeddings in the emotion versus neutral contrasts for all regions (all corrected ps $< .01$, except

Table 2. Results Showing Associations Between Brain Activation Embeddings and Emotional n-Back Task Performance

		Partial Spearman's ρ	95% CI	Partial Spearman's ρ	95% CI
Embedding Type	ROI	2-Back vs. 0-Back		Emotion vs. Neutral Face	
Voxel	Amygdala	0.164	0.122 to 0.193	0.003	−0.030 to 0.029
	Hippocampus	0.154	0.132 to 0.177	0.000	−0.030 to 0.029
	Dorsal attention network	0.157	0.130 to 0.183	0.000	−0.035 to 0.029
	Ventral attention network	0.199	0.176 to 0.220	0.006	−0.018 to 0.027
	Frontoparietal network	0.176	0.152 to 0.196	−0.008	−0.031 to 0.019
PHATE	Amygdala	0.217	0.189 to 0.240	0.113	0.086 to 0.143
	Hippocampus	0.259	0.232 to 0.285	0.114	0.090 to 0.138
	Dorsal attention network	0.540	0.518 to 0.566	0.141	0.118 to 0.168
	Ventral attention network	0.523	0.507 to 0.539	0.150	0.128 to 0.168
	Frontoparietal network	0.534	0.512 to 0.556	0.137	0.111 to 0.160

ROI, region of interest.

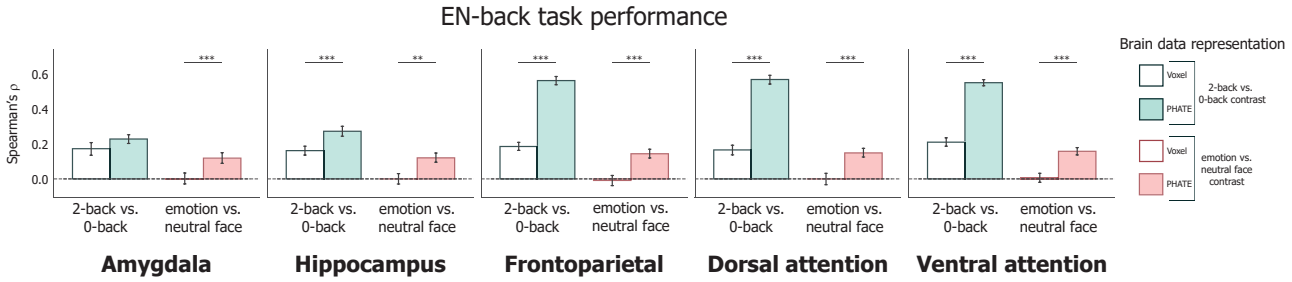

Figure 2. Associations of brain data with in-scanner emotional n-back (EN-back) task performance. Multiple linear regression was used to measure the association between brain data representations and EN-back task performance and scored as the partial Spearman correlation (ρ) between true and regression-predicted performance for held-out participants (20 cross-validation folds). Bars represent the average ρ across cross-validation test folds. Pairwise permutation tests were used to compute p values for the difference in ρ for each cross-validation fold across representations of brain data (i.e., voxel-resolution data in white and corresponding PHATE embedding filled in). Left bar set in each graph represents the 2-back vs. 0-back contrast; right bar set in each graph represents the emotion vs. neutral contrast. Error bars represent the 95% CI of the mean ρ across 20 cross-validation folds. $^{**}p < .01$, $^{***}p < .001$.

amygdala E-PHATE vs. voxel $p < .05$). The magnitude of E-PHATE's performance with emotion versus neutral contrast was comparably strong to E-PHATE's performance with the 2-back versus 0-back contrast, with moderate-to-large effect sizes (all ρs > 0.12) (Figure 3A and Table 2; Supplemental Data 2 in Supplement 3).

CBCL total problems can be broken into 2 main broadband scales: externalizing and internalizing problems. To examine whether the significant relationships between E-PHATE embeddings and total problems were driven specifically by associations with externalizing or internalizing problems, we repeated the analyses above to predict t scores for each scale independently. E-PHATE embeddings were more strongly related to externalizing problems than the voxelwise data or PHATE embeddings, across both task contrasts in all regions (pairwise comparison between E-PHATE and PHATE and E-PHATE and voxel for each region; all corrected ps $< .01$, except frontoparietal E-PHATE vs. PHATE $p < .05$) (Figure 3B and Table 3). As with the total problem score, E-PHATE's relationship to externalizing problems showed moderate-to-large effect sizes across regions and contrasts. The magnitude of associations with internalizing problems ($\rho = 0.076$–0.111 across regions and contrasts) was moderate and lower than the larger effect sizes observed for externalizing problems but still significant (95% CIs did not contain 0) (Table 3). This was not the case for voxelwise or most of the PHATE embeddings, except for the PHATE embedding of 2-back versus 0-back activation in the ventral attention network, which showed a small but significant effect ($\rho = 0.056$, 95% CI, 0.034–0.089) (Figure 3C). Supplemental analyses (Figure S2 in Supplement 1 and Supplemental Data 1 in Supplement 2) replicated this pattern across the externalizing and internalizing subscale symptoms: E-PHATE embeddings were significantly associated with all subscales with moderate-to-large effect sizes, with the strength of the association depending upon the subscale, fMRI contrast, and ROI.

Benchmarking and robustness analyses showed that 1) when comparing E-PHATE to other dimensionality reduction methods, E-PHATE embeddings significantly outperformed PCA (43,44) and UMAP (45,46) (Supplemental Methods S7 and Figure S3 in Supplement 1); 2) the increased sensitivity of E-PHATE was attributable to added information specifically about the environment as opposed to an increase in the quantity of data about each participant (E-PHATE control) (Supplemental Methods S8 and Figure S4 in Supplement 1); 3) the increased sensitivity of E-PHATE was attributable to the nonlinear combination of brain and environment (PHATE + environment); and 4) E-PHATE matrices built solely based on neighborhood disadvantage (area deprivation index) or family conflict (47,48) improved associations compared with no environmental information, but neither afforded as great of an improvement as the 5-feature environment view (Supplemental Methods S9 and Figure S4 in Supplement 1).

E-PHATE Improved Longitudinal Prediction of Emotional and Behavioral Problems in the Frontoparietal Network

To evaluate whether the signals highlighted by E-PHATE could enhance our ability to detect brain activation relevant for future emotional and behavioral problems, we asked whether the same embedding of brain and environmental factors at baseline (ages 9–10) could predict emotional and behavioral problems 2 years later (at ages 11–12). Using a subset of the original participants ($n = 2{,}371$ with complete data at the 2-year follow-up), we embedded baseline voxelwise brain data and environment measures with E-PHATE. Then, we trained multiple linear regressions to predict CBCL total, externalizing, and internalizing problems at the 2-year follow-up (controlling for the corresponding baseline CBCL score and scanner serial number during test).

The prediction of the CBCL total problems score was only significant (although the effect size was small) from E-PHATE embeddings of ventral attention network emotion versus neutral face activation ($\rho = 0.044$, 95% CI, 0.011–0.085); these problems were not predicted by either voxel or PHATE (Figure 4 and Table 3; Supplemental Data 1 in Supplement 2). Externalizing problems were significantly predicted by E-PHATE embeddings of amygdala activation in the 2-back versus 0-back contrast (moderate effect size; $\rho = 0.060$, 95% CI, 0.023–0.096) and by E-PHATE embeddings of hippocampus and ventral attention network emotion versus neutral face activation (moderate effect sizes;

Table 3. Results Showing Associations Between Brain Activation Embeddings and Emotional and Behavioral Problems

		Baseline		2 Years	
Embedding Type	ROI	Partial Spearman's ρ	95% CI	Partial Spearman's ρ	95% CI
CBCL Total Problems					
		2-Back vs. 0-Back		2-Back vs. 0-Back	
Voxel	Amygdala	0.009	−0.025 to 0.042	−0.013	−0.052 to 0.025
	Hippocampus	0.032	0.005 to 0.063	−0.025	−0.061 to 0.016
	Dorsal attention network	0.027	−0.002 to 0.056	0.039	−0.004 to 0.078
	Ventral attention network	−0.025	−0.051 to −0.001	0.019	−0.02 to 0.056
	Frontoparietal network	−0.012	−0.043 to 0.012	−0.036	−0.074 to 0.005
PHATE	Amygdala	0.036	0.016 to 0.057	−0.007	−0.049 to 0.042
	Hippocampus	0.051	0.016 to 0.088	0.024	−0.027 to 0.070
	Dorsal attention network	0.056	0.031 to 0.081	0.005	−0.033 to 0.045
	Ventral attention network	0.069	0.044 to 0.104	−0.006	−0.05 to 0.036
	Frontoparietal network	0.059	0.028 to 0.091	−0.016	−0.059 to 0.023
E-PHATE	Amygdala	0.145	0.128 to 0.163	0.031	−0.007 to 0.07
	Hippocampus	0.152	0.130 to 0.174	0.034	−0.009 to 0.084
	Dorsal attention network	0.143	0.120 to 0.178	0.009	−0.035 to 0.054
	Ventral attention network	0.151	0.126 to 0.181	0.022	0.02 to 0.064
	Frontoparietal network	0.144	0.120 to 0.178	0.043	0.001 to 0.089
		Emotion vs. Neutral Face		Emotion vs. Neutral Face	
Voxel	Amygdala	0.032	−0.007 to 0.043	−0.024	−0.06 to 0.011
	Hippocampus	0.006	−0.029 to 0.034	0.02	−0.017 to 0.056
	Dorsal attention network	−0.017	0.000 to 0.042	−0.052	−0.092 to −0.013
	Ventral attention network	−0.009	−0.033 to 0.006	0.034	−0.013 to 0.078
	Frontoparietal network	−0.016	−0.026 to 0.032	−0.016	−0.061 to 0.017
PHATE	Amygdala	0.017	−0.006 to 0.043	−0.007	−0.047 to 0.032
	Hippocampus	−0.004	−0.03 to 0.028	−0.042	−0.083 to 0.003
	Dorsal attention network	0.017	−0.002 to 0.037	−0.001	−0.052 to 0.047
	Ventral attention network	0.019	−0.015 to 0.051	0.002	−0.033 to 0.05
	Frontoparietal network	−0.013	−0.035 to 0.009	−0.031	−0.069 to 0.005
E-PHATE	Amygdala	0.126	0.101 to 0.151	0.039	−0.011 to 0.089
	Hippocampus	0.149	0.114 to 0.183	0.018	−0.023 to 0.065
	Dorsal attention network	0.124	0.099 to 0.152	0.029	−0.006 to 0.071
	Ventral attention network	0.141	0.111 to 0.17	0.044	0.011 to 0.085
	Frontoparietal network	0.134	0.111 to 0.17	0.032	−0.01 to 0.077
CBCL Externalizing Problems					
		2-Back vs. 0-Back		2-Back vs. 0-Back	
Voxel	Amygdala	0.002	−0.026 to 0.026	−0.023	−0.063 to −0.013
	Hippocampus	0.012	−0.015 to 0.036	0.005	−0.049 to 0.052
	Dorsal attention network	0.028	0.001 to 0.062	0.018	−0.023 to 0.053
	Ventral attention network	−0.039	−0.071 to −0.005	0.009	−0.04 to 0.056
	Frontoparietal network	−0.025	−0.051 to 0.003	−0.05	−0.080 to −0.017
PHATE	Amygdala	0.001	−0.017 to 0.024	0.015	−0.025 to 0.060
	Hippocampus	0.039	0.005 to 0.074	0.011	−0.025 to 0.048
	Dorsal attention network	0.031	0.010 to 0.051	−0.016	−0.051 to 0.019
	Ventral attention network	0.051	0.026 to 0.073	−0.003	−0.047 to 0.048
	Frontoparietal network	0.064	0.039 to 0.099	−0.057	−0.093 to −0.021
E-PHATE	Amygdala	0.121	0.094 to 0.145	0.060	0.023 to 0.096
	Hippocampus	0.141	0.118 to 0.166	0.041	−0.006 to 0.079
	Dorsal attention network	0.142	0.117 to 0.173	0.034	−0.01 to 0.074
	Ventral attention network	0.144	0.121 to 0.172	0.024	−0.018 to 0.059
	Frontoparietal network	0.134	0.107 to 0.175	0.035	−0.008 to 0.076

Table 3. Continued

Embedding Type	ROI	Baseline		2 Years	
		Partial Spearman's ρ	95% CI	Partial Spearman's ρ	95% CI
		Emotion vs. Neutral Face		Emotion vs. Neutral Face	
Voxel	Amygdala	0.006	−0.026 to 0.043	0.010	−0.028 to 0.049
	Hippocampus	−0.025	−0.059 to 0.006	0.008	−0.032 to 0.050
	Dorsal attention network	−0.012	−0.036 to 0.011	−0.028	−0.074 to 0.016
	Ventral attention network	−0.006	−0.031 to 0.020	0.015	−0.029 to 0.056
	Frontoparietal network	−0.008	−0.032 to 0.015	0.002	−0.045 to 0.051
PHATE	Amygdala	0.014	−0.009 to 0.037	−0.009	−0.043 to 0.028
	Hippocampus	0.010	−0.014 to 0.032	0.001	−0.038 to 0.038
	Dorsal attention network	0.043	0.022 to 0.066	0.003	−0.043 to 0.05
	Ventral attention network	0.048	0.014 to 0.081	0.035	−0.011 to 0.077
	Frontoparietal network	−0.013	−0.034 to 0.004	−0.024	−0.059 to 0.014
E-PHATE	Amygdala	0.131	0.108 to 0.153	0.044	−0.002 to 0.085
	Hippocampus	0.134	0.098 to 0.169	0.053	0.006 to 0.096
	Dorsal attention network	0.131	0.105 to 0.159	0.022	−0.019 to 0.062
	Ventral attention network	0.142	0.115 to 0.174	0.053	0.012 to 0.090
	Frontoparietal network	0.139	0.114 to 0.174	0.031	−0.015 to 0.079
CBCL Internalizing Problems					
		2-Back vs. 0-Back		2-Back vs. 0-Back	
Voxel	Amygdala	0.007	−0.025 to 0.038	0.006	−0.038 to 0.048
	Hippocampus	0.014	−0.026 to 0.040	−0.010	−0.052 to 0.032
	Dorsal attention network	−0.001	−0.027 to 0.028	0.010	−0.033 to −0.001
	Ventral attention network	−0.019	−0.046 to 0.004	0.007	−0.051 to 0.044
	Frontoparietal network	−0.002	−0.032 to 0.026	−0.003	−0.030 to 0.039
PHATE	Amygdala	0.017	−0.002 to 0.039	−0.017	−0.050 to 0.022
	Hippocampus	0.022	−0.016 to 0.063	0.016	−0.023 to 0.057
	Dorsal attention network	0.022	−0.012 to 0.051	−0.034	−0.071 to 0.012
	Ventral attention network	0.056	0.034 to 0.089	0.031	−0.006 to 0.074
	Frontoparietal network	0.007	−0.012 to 0.031	0.002	−0.044 to 0.041
E-PHATE	Amygdala	0.111	0.090 to 0.133	0.033	−0.005 to 0.074
	Hippocampus	0.105	0.087 to 0.125	0.051	0.01 to 0.10
	Dorsal attention network	0.093	0.073 to 0.133	0.044	0.009 to 0.09
	Ventral attention network	0.104	0.079 to 0.137	0.034	−0.01 to 0.078
	Frontoparietal network	0.107	0.086 to 0.137	0.058	0.022 to 0.104
		Emotion vs. Neutral Face		Emotion vs. Neutral Face	
Voxel	Amygdala	0.013	−0.018 to 0.047	−0.013	−0.053 to 0.023
	Hippocampus	0.004	−0.019 to 0.029	−0.006	−0.041 to 0.03
	Dorsal attention network	−0.018	−0.032 to −0.002	−0.016	−0.063 to 0.021
	Ventral attention network	0.002	−0.022 to 0.033	0.024	−0.018 to 0.069
	Frontoparietal network	−0.014	−0.031 to 0.008	−0.051	−0.084 to −0.022
PHATE	Amygdala	−0.011	−0.036 to 0.016	−0.010	−0.054 to 0.034
	Hippocampus	−0.019	−0.047 to 0.012	−0.004	−0.037 to 0.028
	Dorsal attention network	−0.008	−0.024 to 0.017	−0.035	−0.076 to 0.005
	Ventral attention network	−0.033	−0.064 to −0.005	0.019	−0.023 to 0.074
	Frontoparietal network	−0.013	−0.039 to 0.019	−0.036	−0.07 to −0.003
E-PHATE	Amygdala	0.091	0.062 to 0.118	0.062	0.019 to 0.11
	Hippocampus	0.110	0.079 to 0.139	0.029	−0.017 to 0.076
	Dorsal attention network	0.076	0.052 to 0.107	0.014	−0.027 to 0.067
	Ventral attention network	0.100	0.074 to 0.125	0.054	0.014 to 0.097
	Frontoparietal network	0.091	0.062 to 0.124	0.047	0.004 to 0.092

CBCL, Child Behavior Checklist; E-PHATE, exogenous PHATE; ROI, region of interest.

hippocampus: ρ = 0.053, 95% CI, 0.006–0.096; ventral attention network: ρ = 0.053, 95% CI, 0.012–0.090). Internalizing problems were more strongly predicted longitudinally than externalizing or total problems, with small-to-moderate effect sizes, again only using E-PHATE embeddings (significant in the amygdala, frontoparietal, and attention networks for the emotion vs. neutral face contrast and in the hippocampus, frontoparietal, and dorsal attention networks for the 2-back vs. 0-back contrast) (Figure 4A and Supplemental Data 2 in Supplement 3). Longitudinal predictions were more focused on specific contrasts and regions than cross-sectional associations, possibly suggesting more nuanced, pointed mechanisms related to distinct problems over time.

DISCUSSION

Decades of theories and empirical research indicate that adolescent neurobiology and environmental context interact to shape the development of emotional and behavioral problems. However, prior work has struggled to capture the complexity of this interplay. Here, we used nonlinear manifold learning to 1) test whether we could improve the basic associations between adolescent neurobiology and individual differences in cognition/mental health (21,49) and 2) model brain-environment interactions in a way that reflects their nonlinear multidimensionality and relation to mental health outcomes. Across a large, sociodemographically diverse sample of U.S. adolescents, PHATE embeddings enhanced the association of fMRI task activation in multiple brain regions and networks with individual differences in cognitive processing. Yet, standard PHATE embeddings did not greatly improve associations with emotional and behavioral problems. Using E-PHATE, we vastly improved the detection and prediction of emotional and behavioral problems. Overall, our results demonstrate that manifold learning techniques are well suited to the complexity of multimodal developmental data and have great potential to enhance research on the neurobiology of emotional and behavioral problems in adolescents.

A major goal of developmental science is to characterize the interplay between adolescents and their broader environments to identify early markers of risk and novel targets for intervention (3,4,14,15,28,29). This work offers a substantial methodological advance toward that goal through the development of E-PHATE. E-PHATE offers researchers a data-driven method for capturing the nonlinear interactions between biological and environmental factors, in contrast with prior univariate approaches that have modeled these interactions as a simple product of 2 variables (50) or multivariate approaches that have been limited in combining brain and environment and have focused more on optimizing the brain

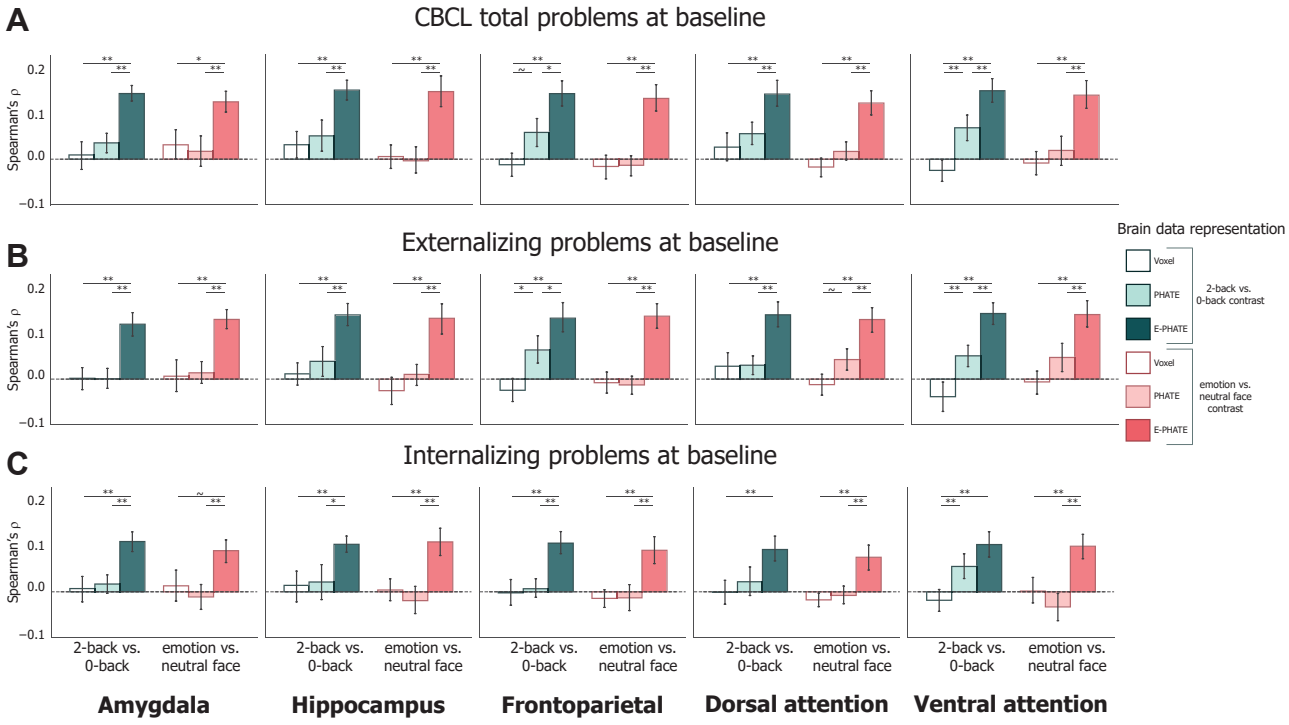

Figure 3. Cross-sectional associations of brain data with mental health problems. Multiple linear regression was used to measure the association between brain data representations and Child Behavior Checklist (CBCL) problem scores and scored as the partial Spearman correlation (ρ) between true and regression-predicted CBCL scores for held-out participants' data (20 cross-validation folds). Bars represent the average ρ across the 20-cross-validation test folds. Pairwise permutation tests were used to compute p values for the difference in ρ for each cross-validation fold across representations of brain data (i.e., voxel, PHATE, exogenous PHATE [E-PHATE]). Both CBCL scores and brain/environmental data were collected at baseline. The left bar set in each graph represents the 2-back vs. 0-back contrast; the right bar set in each graph represents the emotion vs. neutral contrast. **(A)** Total problems, **(B)** externalizing problems, and **(C)** internalizing problems. Error bars represent the 95% CI of the mean ρ across 20 cross-validation folds. $\sim p < .1$, $^{*}p < .05$, $^{**}p < .01$.

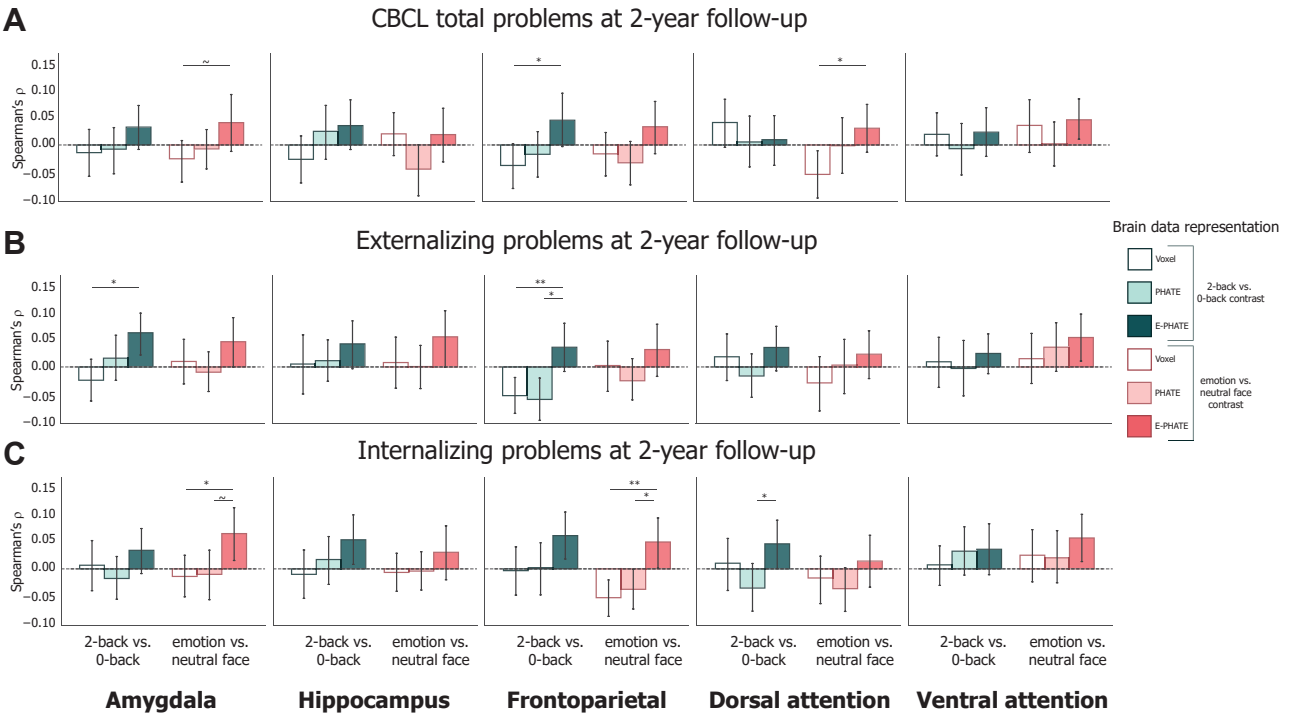

Figure 4. Longitudinal prediction of emotional and behavioral problems. Multiple linear regression analysis was used to predict Child Behavior Checklist (CBCL) problem scores at the 2-year follow-up from baseline brain/environmental data representations and scored as the partial Spearman correlation (ρ) between true and regression-predicted CBCL scores for held-out participants' data (20 cross-validation folds). Bars represent the average ρ across test folds. Pairwise permutation tests were used to compute p values for the difference in ρ across representations of brain data (i.e., voxel, PHATE, exogenous PHATE [E-PHATE]). The left bar set in each graph represents the 2-back vs. 0-back contrast; the right bar set in each graph represents the emotion vs. neutral contrast. **(A)** Total problems, **(B)** externalizing problems, and **(C)** internalizing problems. Error bars represent the 95% CI of the mean ρ across 20 cross-validation folds. $\sim p < .1$, $^{*}p < .05$, $^{**}p < .01$.

data for prediction (18,19). By incorporating exogenous information about adolescents' neighborhoods and families as essential data that provide structure to the brain activation manifold, E-PHATE improved associations between brain function and emotional and behavioral problems. Previous studies have questioned the reliability of developmental, task-based fMRI (49) and of empirical support linking specific ROIs (e.g., amygdala) to emotional and behavioral problems in youths (51–53); however, E-PHATE highlighted signals relevant for understanding emotional and behavioral problems in every ROI and network that we examined across both contrasts. These results demonstrate that efforts to elucidate relationships between adolescent brain function and emotional and behavioral problems may be stifled if researchers fail to consider the broader context in which the brain develops (29,54,55).

Beyond the applications in developmental neuroscience and clinical psychology outlined above, E-PHATE shows promise for a variety of big-data challenges. E-PHATE addresses a key limitation of many manifold learning methods (e.g., PCA, UMAP, or PHATE), which identify latent structure in a purely unsupervised fashion. In other words, the algorithms do not integrate or evaluate the interplay of multiple measurement types into 1 latent structure. Other methods (e.g., partial least squares regression, PCA ridge regression, or canonical correlation analysis) used to study brain-behavior associations from high-dimensional data do so by refining latent components with the direct goal of maximizing the prediction of a certain variable (e.g., CBCL scores) and then testing those components on out-of-sample data (17–19). In contrast, E-PHATE combines different measurements of the same samples into the manifold calculation but does not refine its representation with any specific goal of downstream prediction; thus, it maintains both the benefits of unsupervised manifold geometry discovery and hypothesized structure.

The current work should be viewed in light of a few limitations. First, we focused on specific regions and networks that have previously been related to memory- and emotion processing and mental health (5,6,14,27,35,40,41). However, investigating other brain areas or whole-brain approaches may be relevant for understanding task performance and emotional and behavioral problems. Second, manifold learning algorithms are not able to discern the direction or specific patterns of brain activation that contribute to associations with task performance and emotional and behavioral problems. Nonlinear manifolds are also challenging to faithfully extend to untrained samples (24,56,57), which is an important future direction needed to increase the impact of this work. Third, while E-PHATE could predict emotional and behavioral problems 2 years later, results from the current study are correlational and

cannot be used to infer causation. Considering research that shows bidirectional relationships between the environment and emotional and behavioral problems, future studies should incorporate manifold learning within other longitudinal designs. Fourth, the results in the current article only reflect a snapshot of development. Because the peak onset of emotional and behavioral problems occurs later in adolescence (1), future research investigating a larger developmental window is needed.

Conclusions

We presented evidence for a complex interplay between environments, brain function, and emotional and behavioral problems as uncovered by a new, general-purpose method with interdisciplinary applications. Ultimately, data-driven, interdisciplinary approaches that characterize adolescents' changing neurobiology within the context of their environments may allow us to identify early markers of risk and novel targets for intervention.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Science Foundation Graduate Research Fellowship Program (Grant No. 2139841 [to ELB]) and the National Institutes of Health (Grant No. R21DA057592 [to AB-S]). This article reflects the views of the authors and may not reflect the opinions or views of the National Institutes of Health or ABCD consortium investigators.

We thank Smita Krishnaswamy, Ph.D., for helpful conversations.

A previous version of this article was published as a preprint on bioRxiv: https://doi.org/10.1101/2024.02.29.582854v3.

Data used in the preparation of this article were obtained from the ABCD Study (https://abcdstudy.org), held in the National Institute of Mental Health Data Archive. This is a multisite, longitudinal study designed to recruit more than 10,000 children ages 9 to 10 years and follow them over 10 years into early adulthood. The ABCD Study is supported by the National Institutes of Health and additional federal partners under Grant Nos. U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this report. The ABCD data repository grows and changes over time. The ABCD data used in this report came from https://nda.nih.gov/study.html?id=1299. Our analysis pipeline is available at: https://github.com/ericabusch/manifold_abcd_psychopathology_bpcnni. Stand-alone software package for E-PHATE is available at: https://github.com/ericabusch/EPHATE.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychology, Yale University, New Haven, Connecticut.

ELB and MIC contributed equally to this article.

Address correspondence to Erica L. Busch, M.S., M.Phil., at erica.busch@yale.edu.

Received May 6, 2024; revised Jun 10, 2024; accepted Jul 1, 2024.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2024.07.001.

REFERENCES

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE (2005): Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 62:593–602.
2. Copeland WE, Alaie I, Jonsson U, Shanahan L (2021): Associations of childhood and adolescent depression with adult psychiatric and functional outcomes. J Am Acad Child Adolesc Psychiatry 60:604–611.
3. Sameroff A (2010): A unified theory of development: A dialectic integration of nature and nurture. Child Dev 81:6–22.
4. Bronfenbrenner U, Ceci SJ (1994): Nature-nurture reconceptualized in developmental perspective: A bioecological model. Psychol Rev 101:568–586.
5. Paus T, Keshavan M, Giedd JN (2008): Why do many psychiatric disorders emerge during adolescence? Nat Rev Neurosci 9:947–957.
6. LeDoux JE (2000): Emotion circuits in the brain. Annu Rev Neurosci 23:155–184.
7. Ochsner KN, Gross JJ (2008): Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. Curr Dir Psychol Sci 17:153–158.
8. Conley MI, Hernandez J, Salvati JM, Gee DG, Baskin-Sommers A (2023): The role of perceived threats on mental health, social, and neurocognitive youth outcomes: A multicontextual, person-centered approach. Dev Psychopathol 35:689–710.
9. Fowler PJ, Tompsett CJ, Braciszewski JM, Jacques-Tiura AJ, Baltes BB (2009): Community violence: A meta-analysis on the effect of exposure and mental health outcomes of children and adolescents. Dev Psychopathol 21:227–259.
10. Estrada S, Gee DG, Bozic I, Cinguina M, Joormann J, Baskin-Sommers A (2023): Individual and environmental correlates of childhood maltreatment and exposure to community violence: Utilizing a latent profile and a multilevel meta-analytic approach. Psychol Med 53:189–205.
11. Quon EC, McGrath JJ (2014): Subjective socioeconomic status and adolescent health: A meta-analysis. Health Psychol 33:433–447.
12. Lupien SJ, McEwen BS, Gunnar MR, Heim C (2009): Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat Rev Neurosci 10:434–445.
13. McEwen BS (2012): Brain on stress: How the social environment gets under the skin. Proc Natl Acad Sci U S A 109(suppl 2):17180–17185.
14. Conley MI, Rapuano KM, Benson-Williams C, Rosenberg MD, Watts R, Bell C, *et al.* (2023): Executive network activation moderates the association between neighborhood threats and externalizing behavior in youth. Res Child Adolesc Psychopathol 51:789–803.
15. Weissman DG, Gelardi KL, Conger RD, Robins RW, Hastings PD, Guyer AE (2018): Adolescent externalizing problems: Contributions of community crime exposure and neural function during emotion introspection in Mexican-origin youth. J Res Adolesc 28:551–563.
16. 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.
17. Petrican R, Miles S, Rudd L, Wasiewska W, Graham KS, Lawrence AD (2021): Pubertal timing and functional neurodevelopmental alterations independently mediate the effect of family conflict on adolescent psychopathology. Dev Cogn Neurosci 52:101032.
18. Petrican R, Fornito A (2023): Adolescent neurodevelopment and psychopathology: The interplay between adversity exposure and genetic risk for accelerated brain ageing. Dev Cogn Neurosci 60:101229.
19. Ziegler G, Dahnke R, Winkler AD, Gaser C (2013): Partial least squares correlation of multivariate cognitive abilities and local brain structure in children and adolescents. NeuroImage 82:284–294.
20. Feilong M, Guntupalli JS, Haxby JV (2021): The neural basis of intelligence in fine-grained cortical topographies. eLife 10:e64058.
21. Busch EL, Rapuano KM, Anderson KM, Rosenberg MD, Watts R, Casey BJ, *et al.* (2024): Dissociation of reliability, heritability, and

predictivity in coarse- and fine-scale functional connectomes during development. J Neurosci 44.
22. Moon KR, Stanley JS, Burkhardt D, Van Dijk D, Wolf G, Krishnaswamy S (2018): Manifold learning-based methods for analyzing single-cell RNA-sequencing data. Curr Opin Syst Biol 7:36–46.
23. Moon KR, van Dijk D, Wang Z, Gigante S, Burkhardt DB, Chen WS, *et al.* (2019): Visualizing structure and transitions in high-dimensional biological data. Nat Biotechnol 37:1482–1492.
24. Huang J, Busch E, Wallenstein T, Gerasimiuk M, Benz A, Lajoie G, *et al.* (2022): Learning shared neural manifolds from multi-subject FMRI data. Presented at IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP), August 22–25, Xi'an, China.
25. Rieck B, Yates T, Bock C, Borgwardt K, Wolf G, Turk-Browne N, Krishnaswamy S (2020): Uncovering the topology of time-varying fMRI data using cubical persistence. Adv Neural Inf Process Syst 33:6900–6912.
26. Busch EL, Huang J, Benz A, Wallenstein T, Lajoie G, Wolf G, *et al.* (2023): Multi-view manifold learning of human brain-state trajectories. Nat Comput Sci 3:240–253.
27. Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, *et al.* (2018): The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Dev Cogn Neurosci 32:43–54.
28. Cicchetti D (1993): Developmental psychopathology: Reactions, reflections, projections. Dev Rev 13:471–502.
29. Guyer AE (2020): Adolescent psychopathology: The role of brain-based diatheses, sensitivities, and susceptibilities. Child Dev Perspect 14:104–109.
30. Achenbach T, Ruffle T (2012): The Child Behavior Checklist and related forms for assessing behavioral/emotional problems and competencies. Pediatr Rev 21:265–271.
31. Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, *et al.* (2013): Function in the human connectome: Task-fMRI and individual differences in behavior. NeuroImage 80:169–189.
32. Hautus MJ (1995): Corrections for extreme proportions and their biasing effects on estimated values of d′. Behav Res Methods Instrum Comput 27:46–51.
33. Makowski D (2018): The psycho package: An efficient and publishing-oriented workflow for psychological science. J Open Source Softw 3:470.
34. Hagler DJ Jr, Hatton S, Cornejo MD, Makowski C, Fair DA, Dick AS, *et al.* (2019): Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. Neuroimage 202:116091.
35. Rosenberg MD, Martinez SA, Rapuano KM, Conley MI, Cohen AO, Cornejo MD, *et al.* (2020): Behavioral and neural signatures of working memory in childhood. J Neurosci 40:5090–5104.
36. Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, *et al.* (2011): The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125–1165.
37. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, *et al.* (2018): Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb Cortex 28:3095–3114.
38. Tian Y, Margulies DS, Breakspear M, Zalesky A (2020): Topographic organization of the human subcortex unveiled with functional connectivity gradients. Nat Neurosci 23:1421–1432.
39. McInnes L, Healy J, Melville J (2020): UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv. doi: 1802.03426.
40. Sripada C, Angstadt M, Rutherford S, Taxali A, Shedden K (2020): Toward a "treadmill test" for cognition: Improved prediction of general cognitive ability from the task activated brain. Hum Brain Mapp 41:3186–3197.
41. Satterthwaite TD, Wolf DH, Erus G, Ruparel K, Elliott MA, Gennatas ED, *et al.* (2013): Functional maturation of the executive system during adolescence. J Neurosci 33:16249–16261.
42. O'Brien KJ, Barch DM, Kandala S, Karcher NR (2020): Examining specificity of neural correlates of childhood psychotic-like experiences during an emotional n-back task. Biol Psychiatry Cogn Neurosci Neuroimaging 5:580–590.
43. Smith SM, Hyvärinen A, Varoquaux G, Miller KL, Beckmann CF (2014): Group-PCA for very large fMRI datasets. NeuroImage 101:738–749.
44. Viviani R, Grön G, Spitzer M (2005): Functional principal component analysis of fMRI data. Hum Brain Mapp 24:109–129.
45. Gonzalez-Castillo J, Fernandez IS, Lam KC, Handwerker DA, Pereira F, Bandettini PA (2023): Manifold learning for fMRI time-varying functional connectivity. Front Hum Neurosci 17:1134012.
46. Casanova R, Lyday RG, Bahrami M, Burdette JH, Simpson SL, Laurienti PJ (2021): Embedding functional brain networks in low dimensional spaces using manifold learning techniques. Front Neuroinform 15:740143.
47. Zucker RA, Gonzalez R, Feldstein Ewing SW, Paulus MP, Arroyo J, Fuligni A, *et al.* (2018): Assessment of culture and environment in the Adolescent Brain and Cognitive Development Study: Rationale, description of measures, and early data. Dev Cogn Neurosci 32:107–120.
48. Fan CC, Marshall A, Smolker H, Gonzalez MR, Tapert SF, Barch DM, *et al.* (2021): Adolescent Brain Cognitive Development (ABCD) study Linked External Data (LED): Protocol and practices for geocoding and assignment of environmental data. Dev Cogn Neurosci 52:101030.
49. Kennedy JT, Harms MP, Korucuoglu O, Astafiev SV, Barch DM, Thompson WK, *et al.* (2022): Reliability and stability challenges in ABCD task fMRI data. NeuroImage 252:119046.
50. Tang W, Yu Q, Crits-Christoph P, Tu XM (2009): A new analytic framework for moderation analysis – Moving beyond analytic interactions. J Data Sci 7:313–329.
51. Grogans SE, Fox AS, Shackman AJ (2022): The amygdala and depression: A sober reconsideration. Am J Psychiatry 179:454–457.
52. Deming P, Heilicher M, Koenigs M (2022): How reliable are amygdala findings in psychopathy? A systematic review of MRI studies. Neurosci Biobehav Rev 142:104875.
53. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, *et al.* (2022): Reproducible brain-wide association studies require thousands of individuals. Nature 603:654–660.
54. Viding E, McCrory E, Baskin-Sommers A, Brito SD, Frick P (2024): An 'embedded brain' approach to understanding antisocial behaviour. Trends Cogn Sci 28:159–171.
55. Astle DE, Bassett DS, Viding E (2024): Understanding divergence: Placing developmental neuroscience in its dynamic context. Neurosci Biobehav Rev 157:105539.
56. Duque AF, Morin S, Wolf G, Moon KR (2020): Extendable and invertible manifold learning with geometry regularized autoencoders. Presented at the IEEE International Conference on Big Data (Big Data), December 1–13, Atlanta, GA.
57. Duque AF, Morin S, Wolf G, Moon KR (2023): Geometry regularized autoencoders. IEEE Trans Pattern Anal Mach Intell 45:7381–7394.