Signatures of pubertal brain development and health revealed through domain adapted brain network fusion

Dominik Kraft[1], Dag Alnæs[2, 3], Tobias Kaufmann[1, 2]

[1] Department of Psychiatry and Psychotherapy, Tübingen Center for Mental Health, University of Tübingen, Tübingen, Germany
[2] Norwegian Centre for Mental Disorders Research, University of Oslo and Oslo University Hospital, Oslo, Norway
[3] Kristiania University College, Oslo, Norway

Address for correspondence:

Dominik Kraft, Dr.
Dominik.Kraft@med.uni-tuebingen.de

Tobias Kaufmann, Prof. Dr.
Tobias.Kaufmann@med.uni-tuebingen.de

Department of Psychiatry and Psychotherapy
Tübingen Center for Mental Health
University of Tübingen
Tübingen, Germany

Keywords: Brain development, puberty, mental health, similarity network fusion, diffusion map embedding, domain adaptation

Abstract

Puberty demarks a period of profound brain dynamics that orchestrates changes to a multitude of neuroimaging-derived phenotypes. This poses a dimensionality problem when attempting to chart an individual's brain development on a single scale. Here, we illustrate shifts in subject similarity of imaging data that relate to pubertal maturation and altered mental health, suggesting that dimensional reference spaces of subject similarity render useful to chart brain dynamics in youths.

Introduction

Recent availability of big data in the neurosciences and sparking technical advances have opened doors toward a system level understanding of high-dimensional, multimodal data, integrating information from genetic, behavioral and neuroimaging sources, amongst others[1]. Such deep phenotyping avenues are holding great promise to unravel the complexity and heterogeneity of mental disorders, where a multitude of factors have been identified as contributors to the risk architectures and clinical phenotypes[2–4]. Multimodal big data, however comes with the curse of dimensionality[5] or hurdles regarding how to efficiently and effectively integrate different information sources in biologically meaningful manners[6,7].

Previous research has approached the task of data integration from various angles, from data concatenation to sophisticated modelling[7] such as similarity network fusion (SNF[8]). First application attempts of SNF to common brain disorders have illustrated its potential for deriving insights from heterogeneous populations such as those with psychiatric (e.g.[9]) or neurological (e.g.[6]) disorders. SNF is an unsupervised technique that integrates unique and complementary information from different data sources, thus placing individuals in a comprehensive and biologically informed feature space, which is defined by the similarity between subjects across all data modalities. To achieve this, SNF exploits the covariance between data modalities. Subsequent dimensionality reduction methods such as diffusion map embedding[10] may reveal dominant axes of inter-subject similarity on which subjects can be localized by a single score.

Similar attempts of charting an individual's position on a data continuum have recently shown success in psychiatry, where mapping dimensions of psychopathology can yield advantages over categorical systems[e.g.,11,12].

A key challenge in dimensions that are based on inter-subject similarity is that newly added samples can inevitably result in a change to the overall similarity structure. Consequently, the score that localizes an individual on the dimension is not stable as would be desirable in biomarker utilities, thus marking a disadvantage compared to other data-derived markers such as polygenic risk scores or measures of brain structure. To overcome this, we here propose a machine learning (ML) framework that learns the mapping from raw structural MRI features to the low dimensional embedding score and through supervised domain adaptation allows to transfer this mapping into new datasets without the need to recalculate the fused network. Figure 1A describes the framework schematically. Our approach comes with advantages over modeling fused networks independently for individual datasets and timepoints: First, our ML model establishes a subject similarity reference space in an independent training sample, allowing for robust predictions at an individual subject`s level in unseen data. Second, domain adaptation offers flexibility to adapt the model to other datasets that have unique characteristics, such as repeated measures in a longitudinal design or heterogeneity that is commonly found in patient samples. To this end, we trained our model in the Philadelphia Neurodevelopmental Cohort (PNC)[13] and withheld data from the target datasets that was used for domain adaptation. We validated our approach in an unseen longitudinal sample from the Adolescent Brain Cognitive Development (ABCD) Study[14] and on a clinical population of subjects from the Healthy Brain Network (HBN)[15] sample. We furthermore show that our model can reveal signatures of pubertal brain development and allows to capture biological variance related to emerging psychopathology.

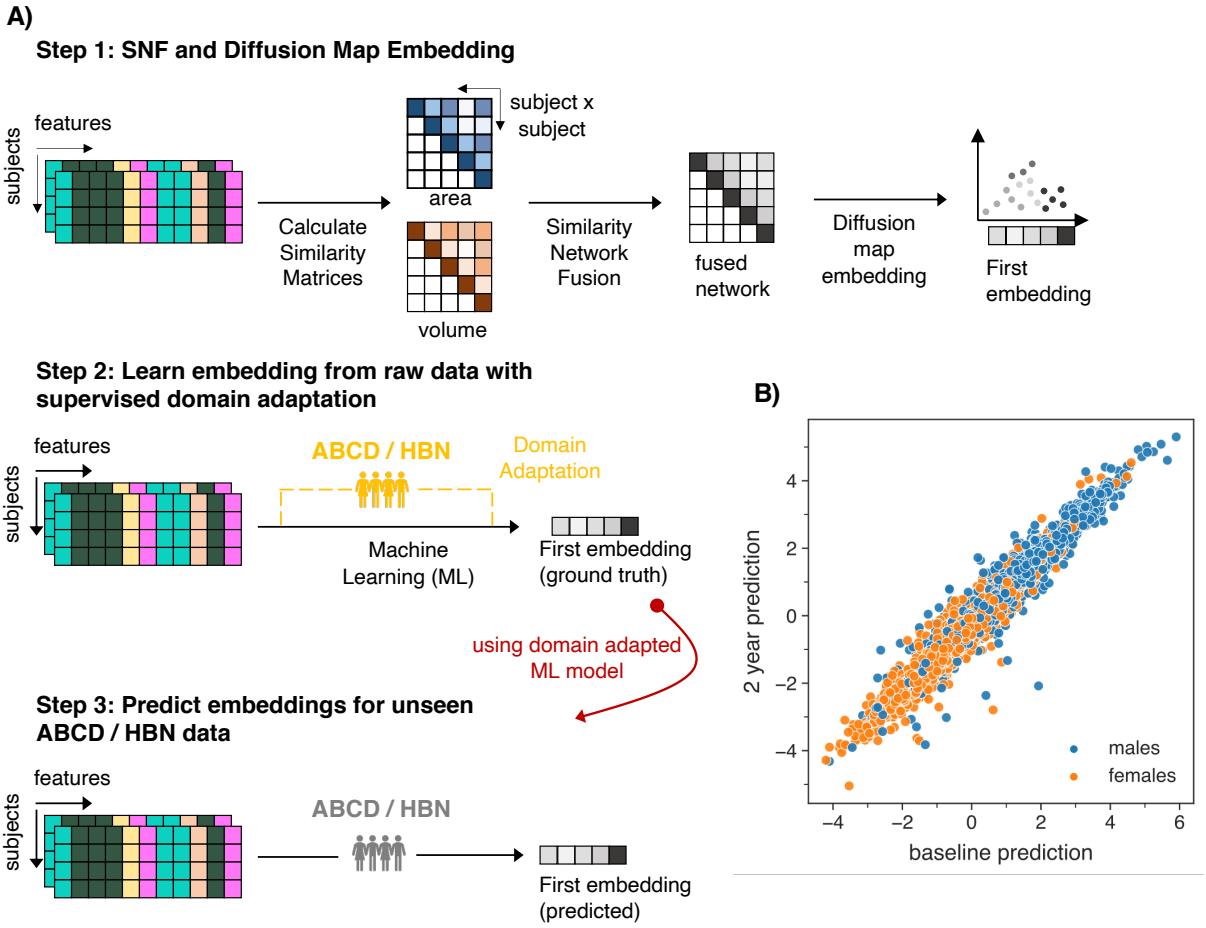

Figure 1. Inferring a reference space using domain adaptation on brain network embeddings. A) Schematic workflow of the prediction framework. Step 1: Similarity network fusion is followed by diffusion map embedding to extract individual subject scores of the first embedding. Step 2: A machine learning model is trained to learn the mapping between raw features and the first embedding score. Using domain adaptation, held-out subsets of the target data (yellow) from ABCD or HBN are added to the training, respectively. Step 3: Such domain adaptation enhances out of sample prediction for unseen (grey) data in both datasets. B) Predicted embeddings for the ABCD baseline (x-axis) and 2-years follow-up (y-axis) data reveal a sex gradient. PNC: Philadelphia Neurodevelopmental Cohort, ABCD: Adolescent Brain Cognitive Development, HBN: Healthy Brain Network.

Results

Model Performance

We applied SNF with subsequent diffusion map embedding to data from N=1594 healthy individuals spanning a wide developmental age range (8 – 21 years, PNC[13].

Akin to other dimensionality reduction approaches, the first embedding captures most variance and was therefore used to build the reference space. We trained a machine learning model with an instance-based domain adaptation procedure (Transfer AdaBoost for Regression)[16] in a combined sample comprising the PNC sample and held-out data from ABCD or HBN to learn the mapping between raw MRI features – specifically cortical area and volume - and the first embedding. This yielded a domain adapted reference model that could be applied to independent data in the ABCD and HBN samples. For the ABCD test sample, we applied the model on baseline and 2-years follow-up data, yielding two predictions per participant. For the cross-sectional HBN sample, the model yielded one prediction per participant. Model performance was calculated by comparing the predicted embeddings in the ABCD and HBN dataset to the 'ground truth' embeddings after performing SNF and diffusion map embedding on the respective test datasets. Our model achieved high performance in unseen data, both for the ABCD and HBN sample (Table 1). Model performance was better in the ABCD sample, which might be driven by the fact, that the sample for domain adaptation in the ABCD dataset was approximately 5x larger than the one used for HBN, allowing for a more efficient shift towards the target distribution. Furthermore, the HBN set comprised data from patients, thus the lower accuracy may to some degree also reflect pathological variance. Moreover, within the ABCD sample, baseline performance was slightly better compared to the follow-up data, since the data used for domain adaptation was also from the baseline study visit. Given successful performance of the model, we proceeded to validating the biological signal in the predictions.

Table 1. Model performance for unseen data in the ABCD and HBN sample.

	RMSE	**MAE**	**R^2**	**r**
$ABCD_{baseline}$	.95	.85	.79	.94
$ABCD_{follow\text{-}up}$	1.02	.91	.78	.94
HBN	1.50	.95	.65	.92

Note: RMSE= root-mean-squared error, MAE= mean absolute error, R^2= coefficient of determination, r= Pearson correlation coefficient.

Biological validation of the model

We validated the biological utility of the predictions in capturing developmental brain dynamics by targeting puberty and mental health as two phenotypes that are closely related to each other and lay off their dynamics during adolescent and therefore are also intertwined with (developmental) brain trajectories[17,18]. We hypothesized that these phenotypes should be related to our embedding score. To assess the models` ability to capture variance cross-sectionally, we first calculated puberty associations for both timepoints and their respective embeddings in the ABCD sample, accounting the statistical model for age and scan site. We observed associations between the average puberty score measured with the Pubertal Development Scale (PDS)[19] and the predicted embedding at both timepoints for the caregiver reports ($baseline_{female}$: b= -.34, $p = 6.88 \times 10^{-16}$, η^2= .02, N= 3344; $baseline_{male}$: b= -.36, $p = 5.53 \times 10^{-10}$, η^2= 0.01, N=3920; $follow\text{-}up_{female}$: b= -.27, $p = 1.94 \times 10^{-15}$, η^2= .03, N= 3316; $follow\text{-}up_{male}$: b= -.17, $p = 1.39 \times 10^{-15}$, η^2= .008, N=3910). In youth reports we observed similar effects although some did not survive Bonferroni correction ($baseline_{female}$: b= -.17, p = .005, η^2= .006, N= 1479; $baseline_{male}$: b= -.06, p = .34, η^2= .0005, N= 2264; $follow\text{-}up_{female}$: b= -.20, $p = 2.66 \times 10^{-9}$, η^2= .02, N= 3271; $follow\text{-}up_{male}$: b= -.14, p = .0003, η^2= .006, N=4056; see Figure 2).

Beyond the cross-sectional associations, the framework allows to apply the model to longitudinal data of the same subjects and investigate change scores between timepoints, as the predicted embedding is modeled with respect to the reference and thus remains stable compared to fused networks derived from individual timepoints, which might introduce additional variance when computing the difference score. Accordingly, we argue that the difference between two predicted embeddings (Δembedding) is capable of tracing brain trajectories and thus may serve as a marker for brain dynamics. Hence, we were particularly interested whether the Δembedding captures biologically meaningful pubertal variance and is thus sensitive to biologically relevant processes shaping the human brain. Consequently, we repeated the linear models with Δembedding as dependent and the ΔPDS scores (i.e., the difference between baseline and 2-years follow up) as independent variable. For age adjustment of this longitudinal analysis, we included the age difference between baseline and the 2 years follow-up (Δage) as a covariate.

Whereas cross-sectional effect sizes were comparable between sexes across both timepoints, change association appeared to be more pronounced in females. For females we observed significant associations for both caregiver (b= -.06, p = 2.37 x 10^{-10}, η^2= .02, N= 3135) and youth report (b= -.08, p = 3.79 x 10^{-11}, η^2= .04, N= 1375) whereas for males, associations did not pass adjustment for multiple comparison (Bonferroni-adjusted α = .05/12 = .004; caregiver: b= .01, p =.19, η^2= .0002, N= 3700; youth: b= .01, p =.36, η^2= .0003, N= 2204; see Figure 2). These effects for females were even more pronounced after controlling for baseline puberty status (caregiver: b= -.08, p= 7.88 x 10^{-20}, η^2= .02, youth: b= -.11, p= 2.38 x 10^{-18}, η^2= .04). It is worth noting that whereas age explained some variance in the embedding scores, significant pubertal effects were always larger than the respective age effects (Supplementary Table 1), supporting that the embedding captures variance relevant to pubertal development beyond age related brain changes.

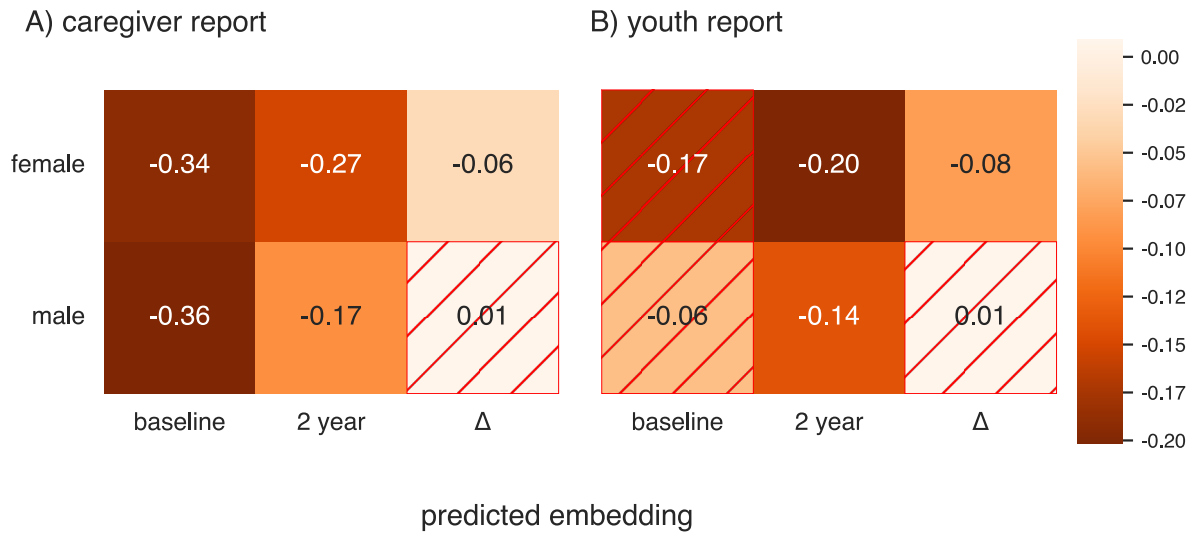

Figure 2: Associations between brain embeddings and puberty, both in cross-sectional and longitudinal data. First two columns in A) and B) refer to associations between predicted embeddings and the respective pubertal score (PDS mean) per timepoint. Δ refers to the association between the Δembedding and the ΔPDS mean score. Annotations refer to effect sizes and hashed cells indicate non-significant results.

Related, we observed that Δembeddings are distributed quite equally across males and females for early pubertal stages, whereas from the 'midpubertal' period onwards distributions start to diverge with respect to earlier developmental stages but also with respect to between group differences (Figure 3). Interestingly, deviations between sexes get even more pronounced with females` menarche, that marks the onset of the late pubertal state.

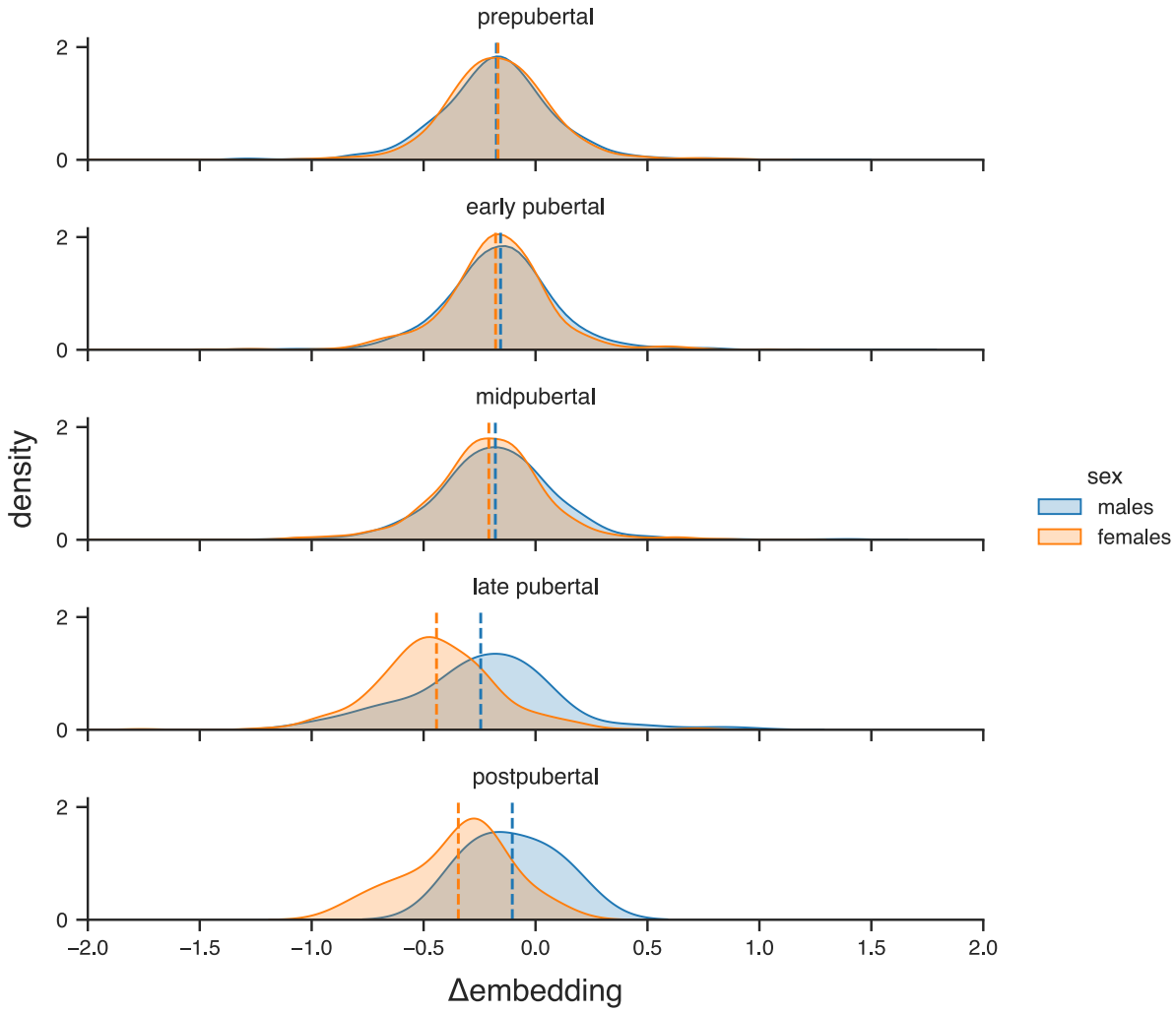

Figure 3: Distribution of Δembedding in the ABCD sample stratified for sex and pubertal categories at one year follow up. Pubertal categories are based on youth report, but caregiver-based categories follow the same pattern (Suppl. Fig. S1). Vertical dashed lines indicate the mean Δembedding per group.

Puberty and adolescence depict a time of cascading changes ranging from biological, emotional to social domains and this phase of transition also constitutes a sensitive and critical period for emerging psychopathology and mental disorders[17,20,21]. Assuming that mental disorders emerge as deviations from a brain 'norm'[22] we argue that our approach of modeling the low dimensional representation anchored to a healthy sample may allow to exploit the resulting reference space (i.e., the embedding) in a normative fashion. To validate this, we tested in a sample of patients drawn from the HBN[15] cohort for associations between the predicted brain embedding score and mental health. For this purpose, we calculated a proxy measure for psychopathology severity, that is the sum of all diagnoses per subject. Participants had between 1 and 10 diagnoses (mean= 2.71, std= 1.6). Using this proxy measure as independent variable together with age and site as covariates, we did not observe a significant effect of psychopathology on the embedding for males (b= .04, p = .08, η^2 = .001, N=1113), but for females the effect survived multiple comparison correction (b= .07, p= .009, η^2 = .009, N=596; Bonferroni-adjusted α = .05/2 = .025). The identified association remained significant when covarying for PDS (females: b= .07, p= .01, η^2 = .010, N=590), yet PDS itself was not significant in this cross-sectional sample, nor were interaction terms between puberty and psychopathology (see Suppl. Table 2 for all effects).

Discussion

The present work illustrates a proof of concept for a new approach that allows to map high dimensional brain imaging data into a low dimensional embedding score which can be then transferred to new datasets by means of domain adaptation and machine learning. By doing so, our framework builds upon similarity network fusion[8] integrating information from different data sources, but does not suffer under the instability of similarity measures and thus can be translated to datasets with unique features such as longitudinal study designs or clinical cohorts without the need to recalculate a fused network in the new sample.

To validate our framework and to test its applicability to other datasets, we trained our model in a sample of healthy subjects spanning a wide age range from the PNC cohort[13] with simultaneous supervised domain adaptation and tested it on two independent validation samples, that is longitudinal data from the ABCD Study[14] and a clinical population of subjects from the HBN sample[15]. Domain adaptation in both datasets was enhanced with independent data that was later not used in the prediction process, such as data from ABCD subjects for whom only baseline data was available and participants in the HBN cohort without clinical diagnoses. Model performance was high for unseen test data in both datasets, confirming the model's ability to generalize to other cohorts. Our approach thus proved useful in two unseen datasets that both displayed unique sample characteristics. We hypothesize that the good model performance also relies on choosing the PNC sample as a source task which stores a rich repertoire of (dis)similarities between healthy participants, from which the domain adaptation procedure for the two new datasets could have been benefitted. However, we consider it important to further investigate the frameworks` boundaries in terms of sample characteristics of the source and target datasets, that is, under which condition the model performance diminishes.

Beyond model building, we aimed at investigating whether the low dimensional embedding scores are sensitive to capture biologically meaningful variance in processes shaping the brain and thus may represent a useful imaging phenotype for (developmental) brain dynamics. Related to work suggesting a close link between pubertal dynamics and neurodevelopment[e.g.,18,23], we observed significant cross-sectional associations between the predicted embedding scores and puberty measures for all models at all timepoints except for baseline data based on youth reports, which might have been biased by the difficulty to rate one´s own pubertal maturation at these early ages. Of note, all analyses were performed stratified for sex, because the embeddings span a sex-gradient, and pubertal timing and trajectories are known to vary between females and males[24]. Moreover, for models in which we observed significant puberty effects, these effects were always larger than the respective age effects, supporting its sensitivity to puberty specific dynamics beyond age.

Since the ABCD study offers an unprecedented resource for granular investigations of child and adolescent brain and pubertal maturation, we leveraged the longitudinal data of the ABCD cohort and investigated whether the Δembedding, that is the difference between the two predicted embedding scores for baseline and the 2-years follow up data, can serve as an additional marker for brain trajectories. Pubertal associations with the Δembedding were significant for females, but not for males. Moreover, it appears that the Δembeddings for both sexes follow a comparable distribution in early pubertal stages, whereas from females´ menarche onwards, both patterns start to deviate from each other. Upcoming releases of the ABCD data may help to further investigate the Δembedding and its ability to capture subtle biological processes like pubertal maturation. With additional longitudinal data one would also expect to have access to more datapoints that represent male participants in later pubertal stages potentially allowing to better disentangle the putative brain trajectories encoded by the Δembedding. However, the significant association of the Δembedding with the pubertal Δ, also when controlling for overall puberty status at baseline and Δage, suggests that our framework can capture longitudinal dynamics of development. While studies on normative brain development generally report overarching brain trajectories across different brain measures[25], recent work by Bottenhorn and colleagues[26] highlight a high degree of intra- and interindividual variability in brain maturation across imaging measures. Identifying and understanding these sources of variance depict an important step towards population-level neuroscience, which however may complicate downstream analyses because of the heterogeneity across regions and imaging measures. Because of its sensitivity to pubertal processes shaping the human brain, we suggest that our approach may help to unify those different sources of variance into a condensed score that does not only serve as a dimension reduction technique but places individuals in a biologically meaningful feature space.

Since puberty is a critical time window for emerging mental disorders[17,20,21], we aimed at additionally exploiting the models predictions as a 'normative' score and tested its association to psychopathology in the HBN sample. In females only, we observed small yet significant effects of psychopathology severity on the embedding score.
By accounting these analyses for age, we ruled out that these associations simply mimic a larger number of diagnoses with increasing age.

When extending the model with pubertal variables, we did not observe additional effects, which may be related to the distribution of age and pubertal stage in the HBN sample. We acknowledge that the sum of diagnoses in the HBN sample rather depicts a coarse measure of psychopathology which could be substituted with more fine-grained quantities, such as hierarchical representations of psychopathology (HiTOP[27]) in future work. These approaches may help to better disentangle associations between the embedding score or Δembedding and different mental health spectra. Since diverse disorders emerge at different timepoints during life[28], leveraging the longitudinal data from ABCD may further help to investigate the markers sensitivity to capture subtle, but biologically meaningful, mental health processes related to brain dynamics. Additionally, dimensional approaches to psychopathology may help to investigate whether the embedding scores are sensitive to capture neuronal traces of early pubertal timing (e.g., early menarche in females[29] and their relationship to internalizing psychopathology[30]). Unfortunately, these kinds of analyses were not feasible in the HBN sample, because of lacking sample size in the respective pubertal category. To further proof the model usefulness as a normative reference space for mental health, future work could investigate how the predictions for sub-clinical samples, such as first-degree relatives of patients with schizophrenia (SZ), may deviate from healthy samples and or patients[31]. Since it is well known that first degree relatives of SZ patients also show brain alterations, albeit with smaller effect sizes[e.g.,32], one could anticipate an embedding distribution of this sample to be located between the healthy and affected group.

Limitations and Future Directions

Potential limitations might stem from the fact that only two imaging modalities, that is brain volume and surface area, were integrated in our framework. The reasoning behind this decision was twofold: First, since brain volume and area follow a comparable normative developmental trajectory from late childhood into late adolescence[25], building similarity networks on both measures may result in robust and non-sparse reference space that allows to better disentangle sex effects, since additional heteroscedasticity of different imaging measures may be mitigated[26].

Secondly, rather pragmatically, the two MRI features were available across all three datasets, which constitutes an important consideration for all future applications of our framework. However, beyond the proof-of-concept of the current study, we nevertheless deem it important to extend our approach with additional (imaging) modalities to tests its generalizability beyond the two imaging features. Furthermore, adding additive data sources may result in a more holistic phenotype representing brain development or dynamics which may help to explain additional variance in behavioral or mental health measures and thus may substantiate the embedding score utility in capturing brain trajectories.

Conclusion

We introduced a novel approach which allows to integrate high dimensional imaging data into a coherent feature space in which subjects can be localized by a single embedding score. We suggest that transferring this mapping to other datasets results in a new imaging phenotype which inherits a sensitivity to capture meaningful and biologically relevant processes shaping human brain dynamics.

Methods

Sample Descriptions

PNC

As source model we used imaging data from the Philadelphia Neurodevelopmental Cohort (PNC), a large-scale cross-sectional study of child and youth between 8- and 21-years age dedicated to study (brain) development. All PNC study procedures were approved by institutional review boards of the University of Pennsylvania and the Children`s Hospital of Philadelphia. All participants or their caregiver provided written informed consent. Data in the PNC sample was acquired from a single site[13]. We included data from N=1594 individuals with available T1-weighted imaging (females= 834, age: *M*= 14.95, *SD*= 3.69). We used brain area and volume of 68 cortical brain regions matching the Desikan-Killiany atlas[33] estimated from T1 MRI images using FreeSurfer (version 7.1.1)[34].

ABCD

The Adolescent Brain Cognitive Development (ABCD) Study is a 10-year longitudinal study of children recruited at age 9 to 10 aiming at characterizing brain developmental trajectories. Overall ~11.000 children were recruited across 21 different sites in the United States[14]. Study procedures have been approved by either the local site Institutional Review Board (IRB) or by local IRB reliance agreements with the central IRB at the University of California. All participants and their parents provided written informed consent. Data for the current study was obtained from ABCD release 4.0 utilizing phenotypic and imaging data from the baseline and 2-years follow up study visit. Preprocessed imaging data from the Desikan-Killiany atlas (68 regions)[33] were downloaded from the NIMH data archive. Since we were interested in the longitudinal data, we included only children having MRI data from both baseline and 2-years follow up visit (N=7776, females= 3587, $age_{baseline}$: *M*= 9.90, *SD*= .62; $age_{follow\text{-}up}$: *M*= 11.90, *SD*= .65).

HBN

The Healthy Brain Network (HBN) is a community sample of children and adolescent (ages 5 – 21) in the New York area aiming at capturing and investigating the heterogeneity in developmental psychopathology and its biological underpinnings[15]. Imaging data was acquired across four different scanning sites and study procedures were approved by the Chesapeake IRB. All participants or their caregiver provided written informed consent. Brain area and volume of 68 cortical brain regions from T1 MRI images were estimated according to the Desikan-Killiany atlas[33] using FreeSurfer (version 7.1.1)[34]. Based on clinical diagnostic information and the presence of a primary diagnosis, we integrated data from N= 2272 (females= 785, age: *M*= 10.43, *SD*= 3.45) participants.

Model Building and Testing

Brain volume and area from the Desikan-Killiany atlas[33] were used to construct fused similarity networks with *snfpy* (version 0.2.2, https://github.com/rmarkello/snfpy). In the following we will briefly describe the SNF workflow but refer the reader to Wang et al.[8] for a more detailed description: First, we generated subject x subject affinity networks for MRI area and volume by converting between-subject (squared euclidean) distances to similarities with a scaled exponential kernel, respectively. Next, SNF iteratively fused each feature affinity matrix resulting into one symmetric similarity matrix integrating information from all data sources. Both previous steps are governed by the hyperparameters K (i.e., the number of neighbors to consider) and μ (i.e., weighting of between subjects' edges) with $K \in [1, 2, \ldots, i]$, $i \in \mathbb{Z}$ and $\mu \in \mathbb{R}^+$. Markello et al.[6] performed a grid-search across 10.000 hyperparameter combinations and reported consistent embeddings across all combinations (r_{mean}= .97), suggesting a neglectable effect of extensive hyperparameter tuning for consecutive analyses aiming at continuous representations. We thus set K= 30 and μ= *0.8* in accordance with the suggested range of values in *snfpy*. The fused matrix is full rank and can then be either subjected to clustering or dimensionality reduction to achieve continuous representation of the data in a low-dimensional space. Since we were interested in the latter, we performed diffusion map embedding on the fused network to derive low-dimensional representations of the imaging data using *BrainSpace* (version 0.1.3)[35].

Diffusion map embedding is a non-linear dimensionality reduction technique that projects the raw data onto dimensions (i.e., embeddings) that encode the primary axes of between-subject similarity. The resulting embeddings are unitless and subjects can be localized according to their inter-subject similarity along these dimensions[6]. Critically, diffusion map embedding has been shown to be sensitive to non-linear relationships and robust against noise perturbations compared to other techniques, such as Principal Component Analyses (PCA)[10,36]. The diffusion time parameter *t* was set to zero to model the most global relationship of the input data[10]. For further analyses the first embedding was used, as it captures the highest variance akin to PCA.

For our machine learning framework we then trained an Elastic Net in scikit-learn (version 1.0.2)[37] to learn the mappings between the raw feature space (i.e., area and volume MRI data, each with shape 1594 x 34 after averaging features across both hemispheres) and the first embedding. Since our goal was to maximize out of sample generalizability, we 1) trained the model with default parameters aiming at minimizing overfitting to the training set and 2) utilized an instance-based supervised domain adaptation (Transfer AdaBoost for Regression; TrAdaBoostR2)[16] implemented in ADAPT (version 0.4.1)[38]. TrAdaBoostR2 combines a source (PNC) and target data set into a single set and performs reverse boosting in which weights of the source instances poorly predicted decrease at each iteration while the ones of the target instances increases, thus shifting the relative importance towards the target set[16]. To avoid data leakage, we used held-out data from the ABCD and HBN: For the ABCD data we used N= 3984 (females= 2027, age: *M*= 9.95, *SD*= .63) children for which only baseline imaging data was available at release 4.0. In the HBN sample we used imaging data from a healthy sample of N=389 (females= 162, age: *M*= 10.45, *SD*= 3.81) for which no primary diagnosis was reported. Of note, for the latter we did pool subjects with the label ‘no diagnoses’ either based on a complete or aborted evaluation. For both datasets we used brain volume and area from the Desikan-Killiany atlas[33]. Since MRI data was acquired on different scanners both for the ABCD and HBN data, we harmonized both the volume and area imaging data individually using *neuroCombat* (version 0.2.12)[39]. Of note, for the ABCD data, batch correction was performed on individual timepoints and separated for train and test set.

After fitting with domain adaptation, we applied the model to unseen test data from the ABCD and HBN, respectively. To quantify the quality of predictions we additionally also performed SNF and diffusion map embedding on the ABCD and HBN test sample and calculated error metrics (MSE; MAE; RMSE) and R^2 and correlation values between the predicted and 'true' first embedding after orthogonal Procrustes alignment with *mapalign* (version 0.3.0, https://github.com/satra/mapalign)[40]. A schematic representation of the workflow is depicted in Figure 1A.

Modelling Puberty

Pubertal development in the ABCD and HBN sample was assessed with the Pubertal Development Scale (PDS) which was designed to resemble the Tanner stages without the need of a physical examination[19,41]. The child's pubertal development is rated on a four-point Likert scale ranging from 'has not begun' to 'completed' with one exception, that is a binary response item regarding females' menarche. Overall, there are general and sex-specific items that are administered with respect to the biological sex, e.g., voice-deepening or breast development. The rating can be conducted by the children or their caregivers, thus reflecting self or other-perceived pubertal maturation. In the ABCD study both children and caregiver report are available for both timepoints[41], whereas in the HBN study only participant responses are available[15]. Individual item scores were used to calculate the average PDS score (PDS_{mean}) in line with procedure described in Herting et al.[42]. For longitudinal associations, we additionally calculated a ΔPDS score as a marker for pubertal maturation, that is the difference between baseline and 2-years follow up PDS score. Moreover, pubertal category scores were derived for males and females. For males, the sum of three items related to pubic and facial hair growth as well as voice deepening was calculated. For females, pubic hair growth and breast development was summed and information about the menarche was additionally incorporated. Eventually, pubertal scores were converted into pubertal categories ranging from prepubertal to post pubertal based on the ABCD conversion scheme. PDS_{mean} scores were also calculated in the HBN sample to test for out-of-sample replicability and generalizability.

500 **Modelling Psychopathology**

501 In the HBN sample each participant and their caregiver underwent an online version
502 of a semi-structured DSM-5 based psychiatric interview (K-SADS)[43] to derive clinical
503 diagnoses. Consensus diagnoses for each participant are made based on the overlap
504 of the child and caregiver interview by a research clinician[15]. We calculated the sum of
505 all consensus diagnoses per subject as a proxy for psychopathology severity.

506

507 **Association Analyses**

508 All association analyses were performed with *statsmodels* (version 0.13.2)[44]. To test
509 for associations between pubertal development and the predicted embeddings in the
510 ABCD study, we implemented linear models for each timepoint (i.e., baseline and 2-
511 year follow up) with the respective embedding as dependent variable (DV) and the
512 PDS_{mean} score as independent variable (IV) with two-sided significance testing. For all
513 associations we additionally report partial-eta-squared (η^2) per predictor of interest.
514 Since we were particularly interested whether the difference between both
515 embeddings (Δembedding) captures biological variance that is associated to brain
516 dynamics, we performed an additional linear model with Δembedding as DV and the
517 ΔPDS_{mean} score as IV. Analyses were stratified for sex and youth and caregiver reports
518 accounting for differences how pubertal development might be perceived[41]. Despite
519 the rather narrow age range at each study visit, age or Δage (i.e., the difference in age
520 between baseline and 2 years follow-up accounting for variance in between-visit
521 durations) was added as a covariate to the linear model in ABCD, to rule out that
522 putative pubertal effects merely represent aging effects. For the ABCD sample the
523 number of observations varies between models as the amount of missing data is
524 different per timepoint and depends on whether the participants themselves or their
525 caregiver provided the data. In the HBN sample, we tested the association between
526 the predicted embedding (DV) and the sum of diagnoses (IV), which we introduced as
527 a proxy for psychopathology severity. Based on the close relationship between puberty
528 and emerging mental disorders, we additionally calculated linear models which
529 included both the summed diagnoses and the PDS_{mean} score as IVs and one model
530 containing an interaction term summed diagnoses: PDS_{mean} score next to the main
531 effects. PDS_{mean} score was based on participant reports.

For the HBN sample the number of observations varies between models as missing data was excluded on a model-by-model bases, i.e., dependent on the IVs of interest. Linear models were stratified for sex and age and site were added as covariates of no interest. All linear models were Bonferroni corrected for multiple comparisons[45].

Data availability

Data incorporated in this work were gathered from various resources (see acknowledgements) and are shared under data use agreements of the respective cohorts.

Code availability

All code used in this manuscript is available on github (https://github.com/dominikkraft/DomAdapt_BrainNetFusion) and builds upon python 3.7.11. Basic data handling relied on *pandas* (version 1.3.5)[46] and *numpy* (version 1.21.5)[47]. Data visualization relied on matplotlib (version 3.5.1)[48] and seaborn (version 0.11.2)[49].

References

1. Huys, Q. J. M., Maia, T. V. & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat. Neurosci.* **19**, 404–413 (2016).
2. Arango, C. *et al.* Risk and protective factors for mental disorders beyond genetics: an evidence-based atlas. *World Psychiatry* **20**, 417–436 (2021).
3. Fernandes, B. S. *et al.* The new field of 'precision psychiatry'. *BMC Med.* **15**, 80 (2017).
4. Prince, M. *et al.* No health without mental health. *The Lancet* **370**, 859–877 (2007).
5. Ganguli, S. & Sompolinsky, H. Compressed Sensing, Sparsity, and Dimensionality in Neuronal Information Processing and Data Analysis. *Annu. Rev. Neurosci.* **35**, 485–508 (2012).
6. Markello, R. D. *et al.* Multimodal phenotypic axes of Parkinson's disease. *Npj Park. Dis.* **7**, 6 (2021).
7. Zitnik, M. *et al.* Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities. *Inf. Fusion* **50**, 71–91 (2019).
8. Wang, B. *et al.* Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods* **11**, 333–337 (2014).
9. Hong, S.-J. *et al.* Toward Neurosubtypes in Autism. *Biol. Psychiatry* **88**, 111–128 (2020).
10. Coifman, R. R. *et al.* Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc. Natl. Acad. Sci.* **102**, 7426–7431 (2005).
11. Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A. & Daw, N. D. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife* **5**, e11305 (2016).
12. Parkes, L. *et al.* Transdiagnostic dimensions of psychopathology explain individuals' unique deviations from normative neurodevelopment in brain structure. *Transl. Psychiatry* **11**, 232 (2021).
13. Satterthwaite, T. D. *et al.* The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage* **124**, 1115–1119 (2016).
14. Casey, B. J. *et al.* The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* **32**, 43–54 (2018).
15. Alexander, L. M. *et al.* An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci. Data* **4**, 170181 (2017).
16. Pardoe, D. & Stone, P. Boosting for Regression Transfer. in *Proceedings of the 27th International Conference on International Conference on Machine Learning* 863–870 (Omnipress, 2010).
17. Pfeifer, J. H. & Allen, N. B. Puberty Initiates Cascading Relationships Between Neurodevelopmental, Social, and Internalizing Processes Across Adolescence. *Biol. Psychiatry* **89**, 99–108 (2021).
18. Vijayakumar, N. *et al.* A longitudinal analysis of puberty-related cortical development. *NeuroImage* **228**, 117684 (2021).
19. Petersen, A. C., Crockett, L., Richards, M. & Boxer, A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *J. Youth Adolesc.* **17**, 117–133 (1988).

20. Dalsgaard, S. *et al.* Incidence Rates and Cumulative Incidences of the Full Spectrum of Diagnosed Mental Disorders in Childhood and Adolescence. *JAMA Psychiatry* **77**, 155 (2020).
21. Paus, T., Keshavan, M. & Giedd, J. N. Why do many psychiatric disorders emerge during adolescence? *Nat. Rev. Neurosci.* **9**, 947–957 (2008).
22. Marquand, A. F., Rezek, I., Buitelaar, J. & Beckmann, C. F. Understanding Heterogeneity in Clinical Cohorts Using Normative Models: Beyond Case-Control Studies. *Biol. Psychiatry* **80**, 552–561 (2016).
23. Holm, M. C. *et al. Linking brain maturation and puberty during early adolescence using longitudinal brain age prediction in the ABCD cohort*. http://medrxiv.org/lookup/doi/10.1101/2022.05.16.22275146 (2022) doi:10.1101/2022.05.16.22275146.
24. Hoyt, L. T., Niu, L., Pachucki, M. C. & Chaku, N. Timing of puberty in boys and girls: Implications for population health. *SSM - Popul. Health* **10**, 100549 (2020).
25. Bethlehem, R. A. I. *et al.* Brain charts for the human lifespan. *Nature* **604**, 525–533 (2022).
26. Bottenhorn, K. L., Cardenas-Iniguez, C., Mills, K. L., Laird, A. R. & Herting, M. M. *Profiling intra- and inter-individual differences in child and adolescent brain development*. http://biorxiv.org/lookup/doi/10.1101/2022.12.19.521089 (2022) doi:10.1101/2022.12.19.521089.
27. Conway, C. C., Forbes, M. K. & South, S. C. A Hierarchical Taxonomy of Psychopathology (HiTOP) Primer for Mental Health Researchers. *Clin. Psychol. Sci.* **10**, 236–258 (2022).
28. Solmi, M. *et al.* Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. *Mol. Psychiatry* **27**, 281–295 (2022).
29. Mendle, J., Ryan, R. M. & McKone, K. M. P. Age at Menarche, Depression, and Antisocial Behavior in Adulthood. *Pediatrics* **141**, e20171703 (2018).
30. Barendse, M. E. A. *et al.* Multimethod assessment of pubertal timing and associations with internalizing psychopathology in early adolescent girls. *J. Psychopathol. Clin. Sci.* **131**, 14–25 (2022).
31. Da Costa, P. F. *et al.* Transformer-based normative modelling for anomaly detection of early schizophrenia. Preprint at http://arxiv.org/abs/2212.04984 (2022).
32. Zwarte, S. M. C. *et al.* Intelligence, educational attainment, and brain structure in those at familial high-risk for schizophrenia or bipolar disorder. *Hum. Brain Mapp.* **43**, 414–430 (2022).
33. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).
34. Fischl, B. *et al.* Whole Brain Segmentation. *Neuron* **33**, 341–355 (2002).
35. Vos de Wael, R. *et al.* BrainSpace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. *Commun. Biol.* **3**, 103 (2020).
36. De la Porte, J., Herbst, B., Hereman, W. & Van Der Walt, S. An introduction to diffusion maps. in *Proceedings of the 19th symposium of the pattern recognition association of South Africa (PRASA 2008), Cape Town, South Africa* 15–25 (2008).
37. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. Preprint at

http://arxiv.org/abs/1201.0490 (2018).
38. de Mathelin, A., Deheeger, F., Richard, G., Mougeot, M. & Vayatis, N. ADAPT : Awesome Domain Adaptation Python Toolbox. Preprint at http://arxiv.org/abs/2107.03049 (2021).
39. Fortin, J.-P. *et al.* Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage* **167**, 104–120 (2018).
40. Langs, G., Golland, P. & Ghosh, S. S. Predicting Activation Across Individuals with Resting-State Functional Connectivity Based Multi-Atlas Label Fusion. in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015* (eds. Navab, N., Hornegger, J., Wells, W. M. & Frangi, A.) vol. 9350 313–320 (Springer International Publishing, 2015).
41. Cheng, T. W. *et al.* A Researcher's Guide to the Measurement and Modeling of Puberty in the ABCD Study® at Baseline. *Front. Endocrinol.* **12**, 608575 (2021).
42. Herting, M. M. *et al.* Correspondence Between Perceived Pubertal Development and Hormone Levels in 9-10 Year-Olds From the Adolescent Brain Cognitive Development Study. *Front. Endocrinol.* **11**, 549928 (2021).
43. Kaufman, J. *et al.* Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 980–988 (1997).
44. Seabold, S. & Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. in 92–96 (2010). doi:10.25080/Majora-92bf1922-011.
45. Bonferroni, C. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni R Ist. Super. Sci. Econ. E Commericiali Firenze* **8**, 3–62 (1936).
46. McKinney, W. Data Structures for Statistical Computing in Python. in 56–61 (2010). doi:10.25080/Majora-92bf1922-00a.
47. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).
48. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
49. Waskom, M. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).

Acknowledgement

The authors (DK, TK) were funded by the Faculty of Medicine, University of Tübingen, and by the Research Council of Norway (# 323961; TK). D.A. is funded by the South-Eastern Norway Regional Health Authority (#2019107, #2020086)

This work was supported by the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) (031A537B, 031A533A, 031A538A, 031A533B, 031A535A, 031A537C, 031A534A, 031A532B). The authors used data from the Philadelphia Neurodevelopmental Cohort (PNC, https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v3.p2, access permission no 29782), the Adolescent Brain Cognitive Development[SM] Study (ABCD, abcdstudy.org), and the Healthy Brain Networks (HBN, data.healthybrainnetwork.org). Support for the collection of the PNC data set was provided by grant #RC2MH089983 awarded to Raquel Gur, MD, PhD, and #RC2MH089924 awarded to Hakon Hakonarson, MD, PhD. ABCD data, held in the NIMH Data Archive (NDA), is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-10 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 award numbers 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/. PNC, HBN, and ABCD consortium investigators designed and implemented the respective studies and/or provided data but did not participate in the analysis or writing of this report. This manuscript reflects the views of the authors and does not necessarily reflect the opinions or views of any other agency, organization, employer or company.

Author Contribution

Dominik Kraft: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Visualization; Writing – original draft; Writing – review, editing, and approval of the paper.
Dag Alnæs: Data curation, Writing – review, editing and approval of the paper.
Tobias Kaufmann: Conceptualization; Project administration; Methodology, Funding acquisition, Writing – original draft; Writing – review, editing, and approval of the paper.

Conflicting Interests

The authors report no conflict of interest.