bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Word: 6,963

Figures: 6

Supplemental Tables: 3

Title: Multimodal parameters of brain structure and function related to sex differences in cognitive performance.

Authors: Ruben C. Gur[1,3,4*], Tyler M. Moore[1], Adon F.G. Rosen[1], Kosha Ruparel[1], Theodore D. Satterthwaite[1], David R. Roalf[1], Efstathios D. Gennatas[1], Warren B. Bilker[1,2], Russell T. Shinohara[2], Allison Port[1], Mark A. Elliott[1,3], Ragini Verma[3], Christos Davatzikos[3], Daniel H. Wolf[1], John A. Detre[4], Raquel E. Gur[1,3,4]

Affiliations:

[1]Brain Behavior Laboratory and the Neuropsychiatry Section, Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, 10th Floor Gates Pavilion, 3400 Spruce Street, Philadelphia, Pennsylvania 19104, USA.

[2]Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine.

[3]Department of Radiology, University of Pennsylvania Perelman School of Medicine

[4]Department of Neurology, University of Pennsylvania Perelman School of Medicine

*Corresponding author Ruben C. Gur, PhD

Brain Behavior Laboratory

10th Floor Gates Pavilion/HUP

3400 Spruce Street

Philadelphia, PA 19104

Tel. (215)6622915

Fax (215)6627903

Email: gur@upenn.edu

Running title: Brain parameters and performance

Keywords: Neuroimaging, multimodal brain parameters, brain performance relation, neurodevelopment, neurocognition

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Cognitive ability is a complex product of brain processes, yet it remains unknown how brain structure and function together lead to individual differences in developing cognitive skills. We performed multimodal neuroimaging in 1601 youths age 8-22 on the same 3-Tesla Magnetic Resonance Imaging scanner with contemporaneous neurocognitive assessment. Across age groups, high performers had larger volumes, greater gray matter density, lower mean diffusivity and lower cerebral blood flow, compared to low performers. These effects, which varied by region, were robust in males and females across age groups, but bigger in females for volume and in adult males for gray matter density and cerebral blood flow. Amplitude of low frequency fluctuations and regional homogeneity values positively related to performance in males but not in adult females for most regions. Regions showing strongest associations with performance include baso-striatal (thalamus), limbic (hippocampus), frontal (orbital and midfrontal) temporal (midtemporal), Parietal (precuneus and superior parietal) and occipital cortex (fusiform and lingual). Cross-validated regularized regressions of brain parameters combined explained ~20% of performance variance. Our cross-modal results indicate that abundance and integrity of neural tissue, as well as the maintenance of low energy metabolism while at rest, combine to optimize cognitive performance in humans.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The relation between brain volume and cognitive performance has received extensive investigation, although its magnitude is debated; estimates of variance in cognitive measures explained by volume range from ~3% to >30% (Witelson, Beresh, & Kigar, 2005; Gignac and Bates, 2017; Nave et al., 2018; Pietschnig et al., 2015). Advanced neuroimaging offers additional parameters of brain structure and function, but few studies linked such parameters to cognitive performance. Ritchie et al. (2015) compared the relationship of volume to performance with more complex neuroanatomic parameters such as cortical thickness and found that volume accounted for the largest share of the variance (around 12%). Ryman et al. (2016) applied graph-theory analyses to volumetric data and reported that in males a latent factor of fronto-parietal gray matter related to general cognitive abilities, while in females it was white matter efficiency and total gray matter volume, with no specificity. Studies with relatively small samples have linked some diffusion tensor imaging (DTI) parameters to performance (e.g., Schmithorst, Wilke, Dardzinski, Holland, 2005; Qiu, Tan, Zhou and Khong, 2008; Genç et al., 2018). A study on a larger sample (N=72) applied graph theory to DTI data and reported sex differences, with females having greater local efficiency (Yan et al., 2011). Studies relating cerebral blood flow (CBF) to performance were done in the elderly, showing that age-related decline in CBF is related to performance decline (Hshieh et al., 2017, Rane et al., 2018). Resting-state functional MRI (rs_fMRI) was examined in relation to cognitive performance in a small sample (Pamplona et al., 2015), and in a larger sample (N=79) by Vakhtin et al. (2014), who examined both resting state and task-activated relations to connectivity. They reported that the regions involved in task related networks included bilateral medial frontal and parietal cortex, right superior frontal lobule, and the right cingulate gyrus. As part of multivariate measures in the Human Connectome Project (Smith et al., 2015), Finn et al. (2015) showed that connectivity profiles predicted “levels

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of fluid intelligence" and that "the same networks that were most discriminating of individuals were also most predictive of cognitive behavior". Indeed, Yoo et al. (2018) reported, across several data sets, that models trained on task data outperformed those trained on rest data in predicting performance. Furthermore, models trained to predict attention performance were effective in predicting not only across samples (Fong et al., 2019), but also reading recall (Jangraw et al., 2018). Notably, Greene, Gao, Scheinost, & Constable (2018) demonstrated, in two large data sets, that predictive models built from task fMRI data outperform models built from resting-state fMRI. Finally, Dubois, Galdi, Paul, & Adolphs (2018) were able to predict up to 20% of variance in general cognitive performance based on resting-state connectivity metrics in a large sample (n=884) from the HCP. However, each of these studies examined individual parameters of either structure or function and there is need for simultaneous examination of neuroanatomic, neurophysiologic and connectivity parameters in relation to cognitive performance. Such simultaneous examination will allow gauging the relative contribution of the anatomic and physiologic parameters to cognitive performance.

We conducted multimodal neuroimaging in a prospective sample of 1601 youths age 8 to 22, all studied on the same high-field (3 Tesla) scanner with contemporaneously obtained measures of cognitive performance, as part of the Philadelphia Neurodevelopmental Cohort (Gur et al., 2012; Calkins et al., 2015). The methods of sample ascertainment and the detailed neuroimaging protocols have been published (Satterthwaite et al., 2014) and are summarized in the Methods. Multimodal neuroimaging yielded regional measures of gray matter (GM) and white matter (WM) volume and GM density (GMD) from T1-weighted scans, mean diffusivity (MD) from diffusion tensor imaging (DTI), resting state cerebral blood flow (CBF) from arterial spin-labeled (ASL) sequences and amplitude of low frequency fluctuations (ALFF) and regional

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

homogeneity (ReHo) measures from rs_fMRI. The methods for image acquisition and processing and for obtaining these brain parameters were detailed in previous publications (Gennatas et al., 2017, Ingalhalikar et al., 2014, Satterthwaite et al., 2014a,b), and are briefly summarized below. The neurocognitive assessment provided measures of accuracy and speed on multiple behavioral domains. Since available literature primarily examined general intellectual functioning, we used as primary cognitive measure the comparable score from the battery, which is a factor score that summarizes accuracy on executive functioning and complex cognition (Moore et al., 2015).

MATERIALS AND METHODS

Participants

Participants for the Philadelphia Neurodevelopmental Cohort (PNC) were recruited from the Children's Hospital of Philadelphia (CHOP) pediatric clinics throughout the Delaware Valley as described in Calkins et al. (2015). A subsample of 1,601 participants (out of the 9498 PNC sample) underwent multimodal neuroimaging as described in Satterthwaite et al. (2014). Of these, 340 were excluded for medical disorders that could affect brain function, such as current use of psychoactive medications, prior inpatient psychiatric treatment, or an incidentally encountered structural brain abnormality. Sample size was further reduced for some modalities upon quality assurance procedures, most for excessive motion. All participants underwent psychiatric assessment (Calkins et al., 2015) and neurocognitive testing (Gur et al., 2012, 2015).

Cognitive measures

Cognitive performance was assessed with the Penn Computerized Neurocognitive Battery (CNB). The CNB consists of 14 tests adapted from tasks applied in functional neuroimaging to

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license

evaluate a range of cognitive domains (Gur et al., 2010, 2012, 2014; Moore et al., 2015; Roalf et al, 2014). These domains include executive control (abstraction and mental flexibility, attention, working memory), episodic memory (verbal, facial, spatial), complex cognition (verbal reasoning, nonverbal reasoning, spatial processing), social cognition (emotion identification, emotion intensity differentiation, age differentiation) and sensorimotor and motor speed. Accuracy and speed for each test were z-transformed. Cognitive performance was summarized by a factor analysis of both speed and accuracy data (Moore et al., 2015), which delineated three accuracy factors corresponding to: 1) executive function and complex reasoning, 2) episodic memory, and 3) social cognition. The first factor was used as the measure of cognitive performance in all analyses.

Neuroimaging

All MRI scans were acquired on a single 3T Siemens TIM Trio wholebody scanner located in the Hospital of the University of Pennsylvania. Signal excitation and reception were obtained using a quadrature body coil for transmit and a 32-channel head coil for receive. Gradient performance was 45 mT/m, with a maximum slew rate of 200 T/m/s.

Structural:

Parameters of brain anatomy were derived from volumetric scans (T1-weighted) and from diffusion tensor imaging.

Volumetric MRI

Brain volumetric imaging was obtained using a magnetization prepared, rapid-acquisition gradient-echo (MPRAGE) sequence. Receive coil shading was reduced by selecting the Siemens

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

prescan normalize option, which is based on a body coil reference scan. Image quality assessment (QA)was performed both by visual inspection and with algorithms to detect artifacts such as related to excessive head motion.

To maximize accuracy, advanced structural image processing, quality assurance and registration procedures were employed for measurement of the cortical volume and gray matter density. Estimation of brain regions used a multi-atlas labeling approach. A set of 24 young adult T1-weighted volumes from the OASIS data set (Marcus et al., 2007) were manually labeled and registered to each subject's T1-weighted volume using the top-performing SyN diffeomorphic registration (Avants et al., 2011a; Klein et al., 2010). These label sets were synthesized into a final parcellation using joint label fusion (Wang et al., 2013). Volume was determined for each parcel using the intersection between the parcel created and prior driven gray matter cortical segmentation from the ANTs cortical thickness pipeline as described below. Density estimates were calculated within each parcel as described below. To avoid registration bias and maximize sensitivity to detect regional effects that can be impacted by registration error, a custom adolescent template and tissue priors were created using data from 140 PNC participants, balanced for age and sex. Structural images were then processed and registered to this custom template using the ANTs cortical thickness pipeline (Tustison et al., 2014). This procedure includes brain extraction, N4 bias field correction (Tustison et al., 2010), Atropos tissue segmentation (Avants et al., 2011b), and SyN diffeomorphic registration method (Avants et al., 2011a; Klein et al., 2010).

Finally, gray matter density was calculated using Atropos (Avants et al., 2011b), with an iterative segmentation procedure that is initialized using 3-class K-means segmentation. This procedure produces both a discrete 3-class hard segmentation as well as a probabilistic gray

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

matter density map (soft segmentation) for each subject. Gray matter density (GMD) was calculated within the intersection of this 3-class segmentation and the subject's volumetric parcellation (Gennatas et al., 2017). Images included in the final analysis passed a rigorous QA procedure as previously detailed (Rosen et al., 2018).

Diffusion (DTI)

Diffusion weighted imaging (DWI) scans for measuring water diffusion were obtained using a twice-refocused spin echo (TRSE) single-shot EPI sequence. The sequence employs a four lobed diffusion encoding gradient scheme combined with a 90-180-180 spin-echo sequence designed to minimize eddy-current artifacts. The sequence consisted of 64 diffusion-weighted directions with b = 1000 s/mm2, and 7 scans with b = 0 s/mm2.

Diffusion data were skull stripped by generating a brain mask for each subject by registering a binary mask of a standard image (FMRIB58_FA) to each subject's brain using FLIRT (Jenkinson et al., 2002). When necessary, manual adjustments were made to this mask. Next, eddy currents and movement were estimated and corrected using FSL's eddy tool (Andersson and Sotiropoulos, 2016; Graham et al., 2016; Roalf et al., 2016). Eddy improves upon FSL's Diffusion Tool Box (Behrens et al., 2003) and eddy correct tool (Andersson and Sotiropoulos, 2016; Graham et al., 2016) by simultaneously modeling the effects of diffusion eddy current and head movement on DTI images, reducing the amount of resampling. The diffusion gradient vectors were rotated to adjust for motion using the 6-parameter motion output generated from eddy. Then, the B0 field map was estimated and distortion correction was applied to the DTI data using FSL's FUGUE (Smith, 2002). Finally, the diffusion tensor was modeled and metrics (FA and MD) were estimated at each voxel using FSL's DTIFIT.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Registration from native space to a template space was completed using DTI-TK (Zhang et al., 2014; Zhang et al., 2006). First, DTI output files from DTIFIT were converted to DTI-TK format. Next, a template was generated from the tensor volumes using 14 representative diffusion data sets that were considered "Excellent" from the PNC sample. One individual from each of the 14 ages (age range 8-21) was randomly selected. These 14 DTI volumes were averaged to create an initial template. Next, data from the 14 subjects were registered to this template in an iterative manner. Unlike standard intensity-based registration algorithms, this process utilizes the full tensor information in an attempt to best align the underlying white matter tracts using iterations of rigid, affine and diffeomorphic registration leading to the generation of a successively refined template. Ultimately, one high-resolution refined template was created and used for registration of the remaining diffusion datasets. All DTI maps were then registered (rigid, affine, diffeomorphic) to the high-resolution study-specific template using DTI-TK. Whole brain analysis was performed using a customized implementation of tract-based spatial statistics (TBSS) (Bach et al., 2014). FA and MD values were computed using a study specific white matter skeleton. Then, standard regions of interest (ROI; ICBM-JHU White Matter Tracts; Harvard-Oxford Atlas) were registered from MNI152 space to the study-specific template using ANTs registration (Avants et al., 2011). Mean diffusion metrics were extracted from these ROIs using FSL's 'fslmeants'. Images included in this final analysis had passed a stringent quality assessment procedure as previously detailed (Roalf et al., 2016).

Functional

Perfusion (ASL)

Brain perfusion was imaged using a pseudo continuous arterial spin labeling (pCASL) sequence (Wu et al., 2007). The sequence used a single-shot spin-echo EPI readout. Parallel acceleration

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(i.e. GRAPPA factor = 2) was used to reduce the minimum achievable echo time. The arterial spin labeling parameters were: label duration = 1500 ms, post-label delay = 1200 ms, labeling plane = 90 mm inferior to the center slice. The sequence alternated between label and control acquisitions for a total of 80 acquired volumes (40 labels and 40 controls), the first being a label.

ASL data were pre-processed using standard tools included with FSL (Jenkinson et al., 2012). Following distortion correction using the B0 map with FUGUE, the first four image pairs were removed, the time series was realigned in MCFLIRT (Jenkinson et al., 2002), the skull was removed with BET (Smith, 2002), and the image was smoothed at 6mm FWHM using SUSAN (Smith and Brady, 1997). CBF was quantified from control-label pairs using ASL Toolbox (Wang et al., 2008). As prior (Satterthwaite et al., 2014), the T1 relaxation parameter was modeled on an age- and sex-specific basis (Wu et al., 2010). This model accounts for the fact that T1 relaxation time differs according to age and sex, and has been shown to enhance the accuracy and reliability of results in developmental samples (Jain et al., 2012). The CBF image was co-registered to the T1 image using boundary-based registration (Greve and Fischl, 2009), and regional CBF values were averaged within each parcel. Subjects included in this analysis had low motion as measured by mean relative framewise displacement less than 2.5 mm.

Resting-state BOLD

Resting-state BOLD scans were acquired with a single-shot, interleaved multi-slice, gradient-echo, echo planar imaging (GE-EPI) sequence. In order to reach steady-state signal levels, the sequence performed two additional dummy scans at the start. The imaging volume was sufficient to cover the entire cerebrum of all subjects, starting superiorly at the apex. In some subjects, the inferior portion of the cerebellum could not be completely included within the imaging volume. The selection of imaging parameters was driven by the goal of achieving whole brain coverage

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license

with acceptable image repetition time (i.e. TR = 3000 ms). A voxel resolution of 3 × 3× 3 mm with 46 slices was the highest obtainable resolution that satisfied those constraints. During the resting-state scan, a fixation cross was displayed as images were acquired. Participants were instructed to stay awake, keep their eyes open, fixate on the displayed crosshair, and remain still. Resting state scan duration was 6.2 min.

Task-free functional images were processed using a top-performing pipeline for removal of motion-related artifact (Ciric et al., 2017). Preprocessing steps included (1) correction for distortions induced by magnetic field inhomogeneities using FSL's FUGUE utility, (2) removal of the 4 initial volumes of each acquisition, (3) realignment of all volumes to a selected reference volume using MCFLIRT (Jenkinson et al., 2002) (4) removal of and interpolation over intensity outliers in each voxel's time series using AFNI's 3DDESPIKE utility, (5) demeaning and removal of any linear or quadratic trends, and (6) co-registration of functional data to the high-resolution structural image using boundary-based registration (Greve and Fischl, 2009). The artifactual variance in the data was modelled using a total of 36 parameters, including the 6 framewise estimates of motion, the mean signal extracted from eroded white matter and cerebrospinal fluid compartments, the mean signal extracted from the entire brain, the derivatives of each of these 9 parameters, and quadratic terms of each of the 9 parameters and their derivatives. Both the BOLD-weighted time series and the artifactual model time series were temporally filtered using a first-order Butterworth filter with a passband between 0.01 and 0.08 Hz.

ReHo: Voxelwise regional homogeneity (ReHo; Zang et al., 2004) is equivalent to Kendall's coefficient of concordance computed over the timeseries in each voxel's local neighborhood. ReHo can thus be used as an estimate of the homogeneity of each neighborhood's

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

activation pattern. Because spatial smoothing intrinsically elevates ReHo estimates by elevating spatial autocorrelation, Kendall's W was computed only on unsmoothed data. Each voxel's neighborhood was defined to include the 26 voxels adjoining its faces, edges, and vertices. The voxelwise homogeneity map was subsequently smoothed using a Gaussian kernel with FWHM of 6mm in SUSAN to improve the signal-to-noise ratio (Smith and Brady, 1997). Finally regional ReHo values were then averaged across the anatomically derived subject specific segmentation. Participants included in this analysis had low motion with mean relative frame wise displacement less than 2.5 mm.

ALFF: Functional connectivity among brain regions is primarily attributable to correlations among low-frequency fluctuations in regional activation patterns. The voxelwise amplitude of low-frequency fluctuations (ALFF; Zang et al., 2007) was computed as the sum (discretised integral) over frequency bins in the low-frequency (0.01-0.08Hz) band of the voxelwise power spectrum, computed using a Fourier transform of the time-domain of the voxelwise signal. ALFF was calculated on data smoothed in SUSAN using a Gaussian-weighted kernel with 6mm FWHM(Smith and Brady, 1997).

Brain and Behavior Associations

Performance prediction

Performance on the CNB (Moore et al., 2015) was then predicted using each aforementioned modality. Subjects included in this analysis passed each modality specific inclusion criteria and received the cognitive battery within a year of the neuroimaging. Age effects were removed from the data by regressing age, the quadratic, and cubic effect of age on each region; finally bilateral

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

regions were averaged across hemispheres. Models predicting performance using all regional values and a global summary metric from each modality were built in a 10 fold cross-validated fashion within each sex. Stratified samples were built using the `createFolds` function from the `caret` (Kuhn et al., 2016) package in R. Within each training fold a ridge regression model was built modeling the performance as predicted by every regional modality estimate and global summary metric using glmnet function in the 'glmnet' (Friedman et al., 2017) package. The optimal lambda was tuned within each training fold by finding the lambda value with the lowest cross-validated error within the training fold. This procedure was performed 1000 times in order to obtain confidence intervals and means of cross-validated R^2 values. Finally, statistical significance was determined by comparing these point estimates to empirical null distributions of cross-validated R^2 values under the null hypothesis (computed by running the models 1000 times on randomly permuted cognitive estimates).

RESULTS

We used two complementary approaches to examine the effects of variability in brain structural and functional parameters and our performance parameter: hypothesis-testing and data-driven, both incorporating procedures to contain type I error. Analyses were conducted using the open source R platform (Version 3.5, R Core Team, 2015).

Hypothesis-testing results

For the hypothesis-testing approach, analyses were conducted at the global and regional levels by fitting generalized estimating equations (GEEs) with unstructured working correlation structure. GEE models are an extension of generalized linear models that estimate dependence among repeated measures by a user-specified working correlation matrix which allows for correlations in the dependent variable over time or across observations. Five nested forms of GEE models

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were fit and evaluated as shown below. The null model (Model 1) evaluated the association of the demographic variables, age and sex, on brain parameters. A second model added the performance term to evaluate the association between performance and brain parameters adjusted for demographic variables (Model 2). To evaluate if association of performance and brain parameters differed by sex or by age, interaction terms were added as shown in Model 3 and Model 4 respectively. Finally, to evaluate if association of performance and brain differed by both age and sex, Model 5 included all main effects and all possible interactions. Model performance was compared using a Wald test. A squared age term was included to capture non-linear effects of age. Models were fit at all anatomical specificity levels.

Model 1: Null model: Sex and Age Associations

$$Y_{ij} = \beta_0 + \beta_1\ (Brain_{ij}) + \beta_2\ (Sex_i) + \beta_3\ (Age_i) + \beta_4\ (Brain_{ij}\ \mathrm{X}\ Sex_i) + \beta_5\ (Brain_{ij}\ \mathrm{X}\ Age_i) + \beta_5\ (Sex_i\ \mathrm{X}\ Age_i) + \beta_6\ (Brain_{ij}\ \mathrm{X}\ Sex_i\ \mathrm{X}\ Age_i) + \varepsilon_{ij}$$

Model 2: Model Associations of Performance and Brain

$$Y_{ij} = \beta_0 + \boldsymbol{\beta_1}\ (\boldsymbol{Performance_i}) + \beta_2\ (Brain_{ij}) + \beta_3\ (Sex_i) + \beta_4\ (Age_i) + \boldsymbol{\beta_4}\ (\boldsymbol{Performance_i}\ \mathbf{X}\ \boldsymbol{Brain_i}) + \beta_5\ (Brain_{ij}\ \mathrm{X}\ Sex_i) + \beta_6\ (Brain_{ij}\ \mathrm{X}\ Age_i) + \beta_7\ (Sex_i\ \mathrm{X}\ Age_i) + \beta_8\ (Brain_{ij}\ \mathrm{X}\ Sex_i\ \mathrm{X}\ Age_i) + \varepsilon_{ij}$$

Model 3: Model with Sex modifying the associations of Performance and Brain

$$Y_{ij} = \beta_0 + \beta_1\ (Performance_i) + \beta_2\ (Brain_{ij}) + \beta_3\ (Sex_i) + \beta_4\ (Age_i) + \beta_4\ (Performance_i\ \mathrm{X}\ Brain_i) + \boldsymbol{\beta_5}\ (\boldsymbol{Performance_i}\ \mathbf{X}\ \boldsymbol{Sex_i}) + \beta_6\ (Brain_{ij}\ \mathrm{X}\ Sex_i) + \beta_7\ (Brain_{ij}\ \mathrm{X}\ Age_i) + \beta_8\ (Sex_i\ \mathrm{X}\ Age_i) + \boldsymbol{\beta_9}\ (\boldsymbol{Performance_i}\ \mathbf{X}\ \boldsymbol{Brain_{ij}}\ \mathbf{X}\ \boldsymbol{Sex_i}) + \beta_{10}\ (Brain_{ij}\ \mathrm{X}\ Sex_i\ \mathrm{X}\ Age_i) + \varepsilon_{ij}$$

Model 4: Model with Age modifying the associations of Performance and Brain

$$Y_{ij} = \beta_0 + \beta_1\ (Performance_i) + \beta_2\ (Brain_{ij}) + \beta_3\ (Sex_i) + \beta_4\ (Age_i) + \beta_4\ (Performance_i\ \mathrm{X}\ Brain_i) + \boldsymbol{\beta_5}\ (\boldsymbol{Performance_i}\ \mathbf{X}\ \boldsymbol{Age_i}) + \beta_6\ (Brain_{ij}\ \mathrm{X}\ Sex_i) + \beta_7\ (Brain_{ij}\ \mathrm{X}\ Age_i) + \beta_8\ (Sex_i\ \mathrm{X}\ Age_i) + \boldsymbol{\beta_9}\ (\boldsymbol{Performance_i}\ \mathbf{X}\ \boldsymbol{Brain_{ij}}\ \mathbf{X}\ \boldsymbol{Age_i}) + \beta_{10}\ (Brain_{ij}\ \mathrm{X}\ Sex_i\ \mathrm{X}\ Age_i) + \varepsilon_{ij}$$

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Model 5: Model with Age and Sex modifying the associations of Performance and Brain

$$Y_{ij} = \beta_0 + \beta_1\ (Performance_i) + \beta_2\ (Brain_{ij}) + \beta_3\ (Sex_i) + \beta_4\ (Age_i) + \beta_5\ (Performance_i \text{ X } Brain_i) + \beta_6\ (Performance_i \text{ X } Sex_i) + \beta_7\ (Performance_i \text{ X } Age_i) + \beta_8\ (Brain_{ij} \text{ X } Sex_i) + \beta_9\ (Brain_{ij} \text{ X } Age_i) + \beta_{10}\ (Sex_i \text{ X } Age_i) + \beta_{11}\ (Performance_i \text{ X} Brain_{ij} \text{ X } Sex_i) + \beta_{12}\ (Performance_i \text{ X} Brain_{ij} \text{ X } Age_i) + \beta_{13}\ (Brain_{ij} \text{ X } Sex_i \text{ X } Age_i) + \boldsymbol{\beta_{14}\ (Performance_i \text{ X } Brain_{ij} \text{ X } Sex_i \text{ X } Age_i)} + \varepsilon_{ij}$$

Global values: The global measures included estimated total brain volume, average whole-brain values of GMD, MD, CBF, ALFF and ReHo. The global measures were scaled within modality prior to fitting the GEE model. This analysis indicated that performance was significantly associated with whole brain global measures across modalities and this association differed by sex and age (GEE, Wald $\chi2 = 59.51$, df=30, p =0.001). Age interactions were examined by dividing the sample into children (ages less than 12), adolescents (ages 12-17) and young adults (ages 18 and older). Performance interactions were examined by dividing the sample into high, middle, and low performance bins based on tertile splits. This interaction (Figure 1, top panel) indicated that while high performers had increased volume and GMD compared to medium and low performance groups, they had lower MD and CBF with no differences in ALFF and ReHo. As can be seen in Figure 1 (bottom panel), effect sizes of differences between performance groups became stronger from childhood to adolescence to adulthood for volume and GMD. For MD and CBF this trend was seen in males, but not in females. For ALFF and ReHo the effect size of increased values in the high performance group reached .5 SDs in the oldest male group, but diminished and even reversed for ALFF in the oldest female group.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

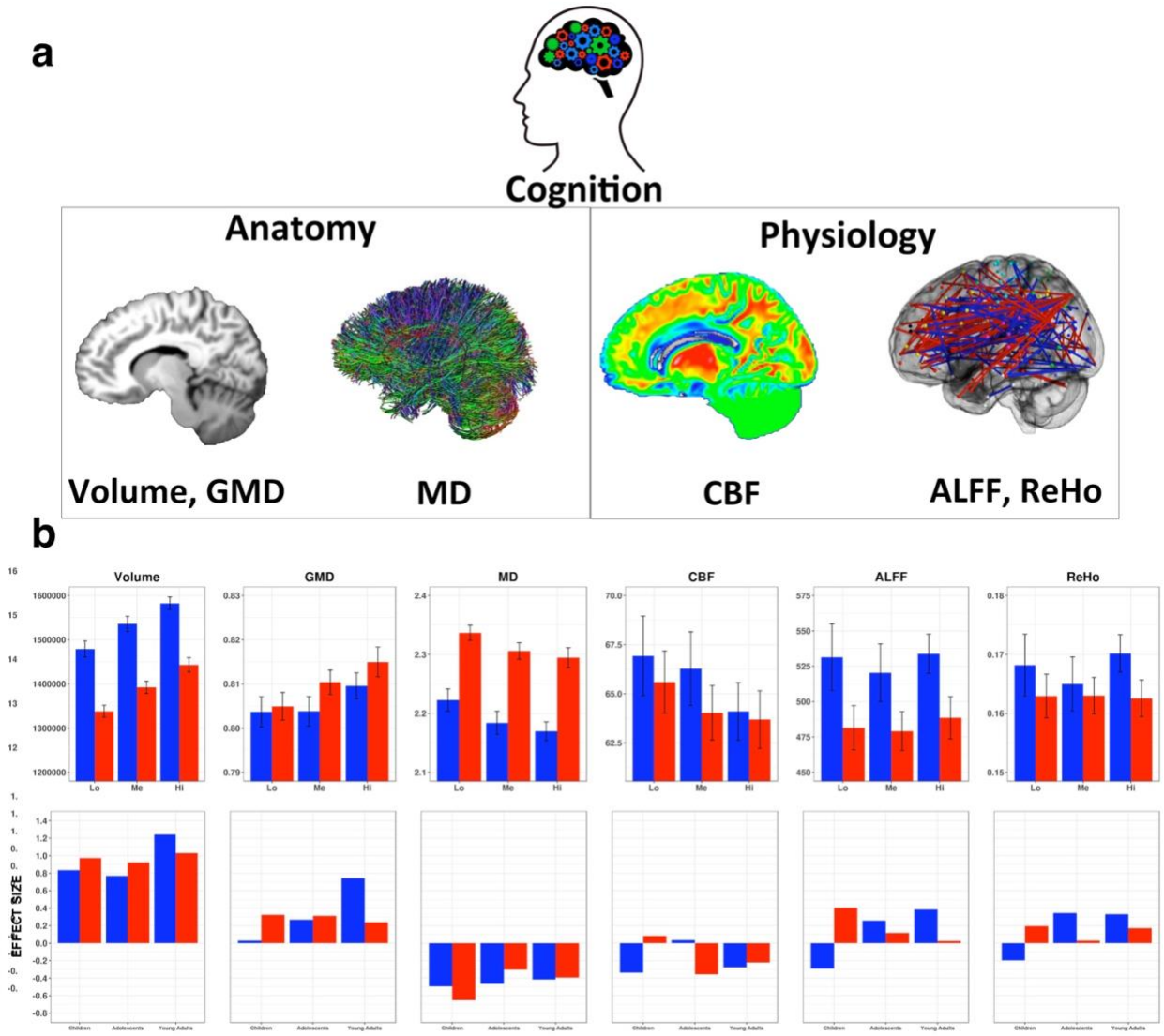
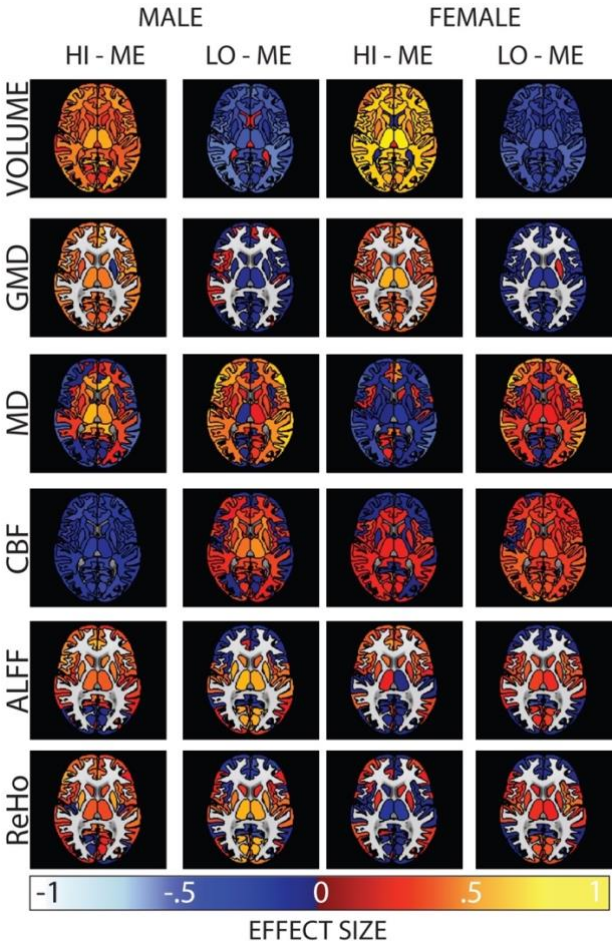


Figure 1. a. Illustration of anatomic and physiologic parameters examined (top panel). b. Global values (±95% CI) for each modality by performance group (tertiles) in males and females. Top panel shows the Performance*Domain interaction displaying means (+/-95% CI) of Low (Lo), Medium (Me) and High (Hi) tertiles of performers on Volume, GMD, MD, CBF, ALFF and ReHo. Bottom panel illustrates the 4-way interaction (Performance*Age*Sex*Domain) by displaying effect sizes (Cohen's D) for the differences between High and Low performers on each parameter.

Regional analyses: To examine regional specificity and compare regional differences of interactions with performance, GEE models were fit within each modality. To minimize type I error given the large number of regions (up to 128 regions per modality), we aggregated them

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

into 8 neuroanatomic Sections: cerebellum, baso-striatal, limbic, frontal, temporal, parietal, occipital, WM. Section volumes were derived using the sum of all regions involved; for all other brain measures a volume-weighted mean of the regions involved was calculated at the subject-level for each brain section. Exploratory analyses to understand neurodevelopmental changes were conducted for those brain sections and in those modalities that showed significant interactions with performance. Significant effects and interactions of performance were elucidated by charting the brain parameter profiles of effect sizes (Cohen D) for the differences between the high and low performance groups. The regional maps of effect sizes contrasting high and low performers to the middle group on each parameter are shown in Figure 2.



bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2. The same axial slice for the six modality contrasts of High vs Medium and Low vs Medium performance group (tertiles) contrast. Effect size units are displayed for high minus medium and low minus medium performance groups in males (left columns) and females (right columns) for volume (top row), GMD (second row), MD (third row), CBF (fourth row), ALFF (fifth row) and ReHo (bottom row). All slices are available as movies here: https://pennbbl.github.io/hiLo/index.html

For volume, there was a significant performance by brain section interaction (GEE, Wald $\chi2 = 191.24$, df=8, $p < 0.0001$) justifying exploratory analyses within each brain section.

The exploratory analyses showed that performance was significantly associated with brain parameters for all brain sections (Supplementary Table S3). Higher brain volume was associated with better performance in specific brain regions in all sections (all Performance x Region interactions highly significant within each section, Table S3). Notably, no higher-order interactions with sex or age were significant, indicating relative uniformity of effects in males and females across youth. To elucidate the performance x region interactions, we therefore examined age-adjusted differences between high and low performance groups. As can be seen in Figure 3 (top panel), effect sizes showing higher volumes in high performance compared to low performance groups ranged from small (.1 to .2 SD difference) to large ($\geq$ .8 SD difference). The largest effect sizes for cerebellum were in the exterior cerebellum and in the cerebellar vermal lobules VII-X (.6 and .8 SDs, respectively), the largest for baso-striatal section was for the thalamus (.6 and .8 SDs for males and females, respectively) and the largest effect size of limbic regions was in the hippocampus (same order of magnitude). Frontal regions showed fairly uniform effect sizes, the largest being for posterior orbital gyrus (>.6 SDs), while temporal regions showed large variability, with the strongest effect sizes for temporal pole, fusiform gyrus, inferior temporal gyrus and mid-temporal gyrus (.6 to .8 SDs). Parietal and occipital lobes also showed variability of effect sizes, the largest effect for parietal was in the precuneus (>.6 SDs), while in occipital the largest effect sizes were in occipital fusiform gyrus, lingual gyrus,

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

inferior occipital gyrus and occipital pole (.7 to .8 SDs). For WM, effects were generally larger for females than males and were the largest for temporal and parietal WM.

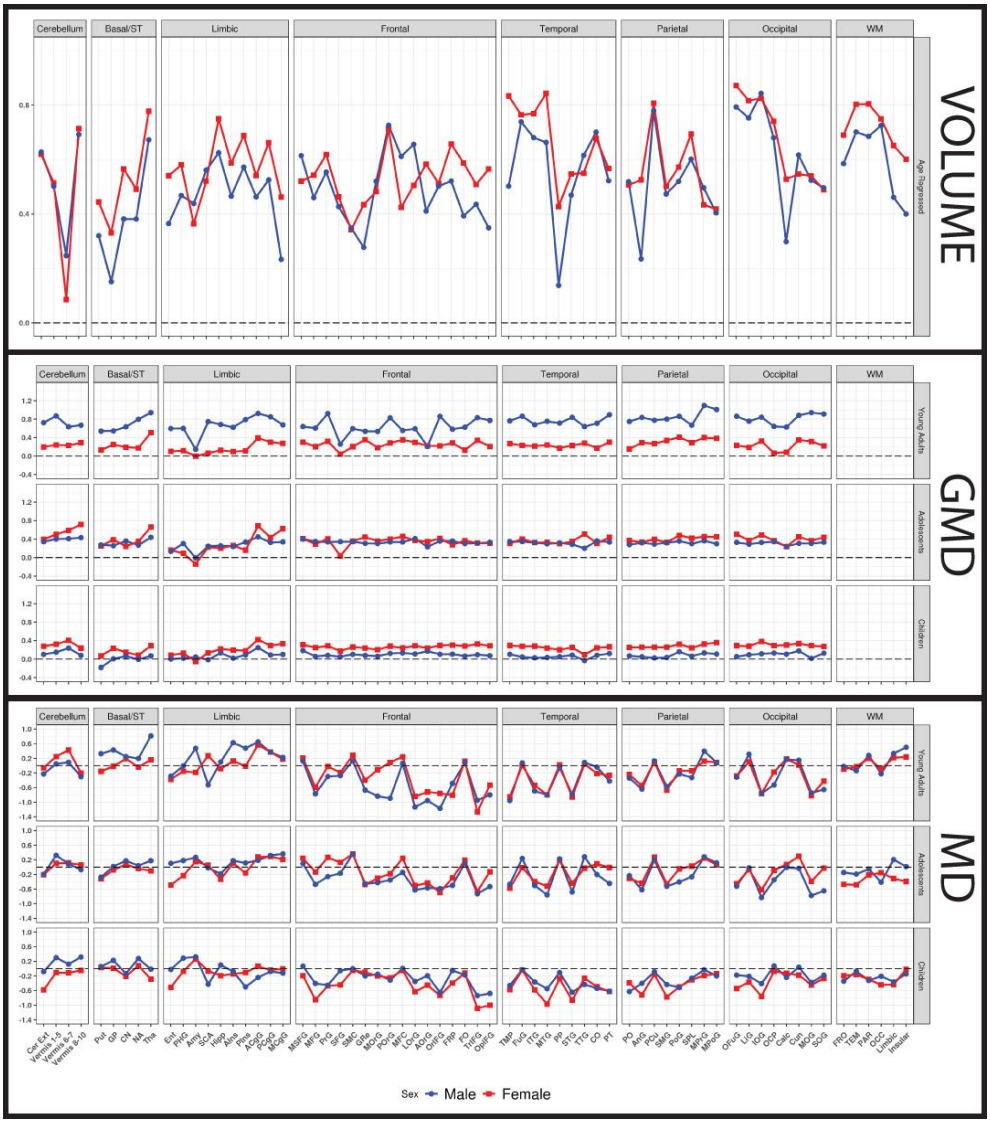


Figure 3. Effect sizes (Cohen D) for Volume, GMD and MD in individual ROIs. Since there was no age interaction for volume, values are shown age-adjusted for the entire sample across the age ranges (top panel). Since there were age interactions for GMD and MD, results are separated by age groups (children, adolescents, young adults) for these parameters (middle and bottom panels). ROI names can be found here: https://pennbbl.github.io/hiLo/oasisLabels/index.html

For GMD, there was a significant performance by brain section interaction that differed by age and sex (GEE, Wald $\chi2$ = 61.15, df=30, p =0.0007) justifying exploratory analyses within each brain section. The exploratory analyses showed that performance was significantly associated with brain parameters for all sections. (Supplementary Table S3). In addition, this

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

association differed significantly by sex and by age within the frontal lobe, and by age for temporal, parietal and occipital lobes (Figure 3, middle panel). High performers had higher GMD across age and sex groups, but in males the effect sizes increased in most regions from childhood to adolescence to young adulthood, while in females they remained generally stable across age bins.

For MD, there was a significant performance by brain section interaction (GEE, Wald $\chi 2 = 108.82$, df=8, $p < 0.0001$) justifying exploratory analyses within each brain section.

The exploratory MD analyses showed significant association of performance and brain region for all sections except for baso-striatal (Table S3, Figure 6). In addition, for all brain sections except WM, the association differed by age and sex. Overall, except for some baso-striatal and limbic ROIs showing the opposite effects, across cortical regions, both male and female high-performing groups had lower MD than the low-performing groups. The effect sizes differentiating high and low performers ranged widely from large (>.8) in some orbital frontal regions to moderate in specific cortical regions. This pattern was similar in males and females, displaying large convergence across effect sizes. The effect sizes increased from childhood to adolescence to young adulthood, especially in frontal regions (Figure 3, bottom panel).

For CBF, there was a significant performance by brain section interaction that also differed by age and sex (GEE, Wald $\chi 2 = 67.15$, df=35, p =0.0.0008) justifying exploratory analyses within each brain section. The exploratory analyses for each brain section showed significant Performance*Region*Age*sex interactions for limbic, frontal and temporal regions, as well as white matter (Table S3). The interactions for these regions (Figure 4, top panel) indicated that effect sizes of lower CBF associated with high performance were stronger for males than females and showed greater increase from childhood to young adulthood.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

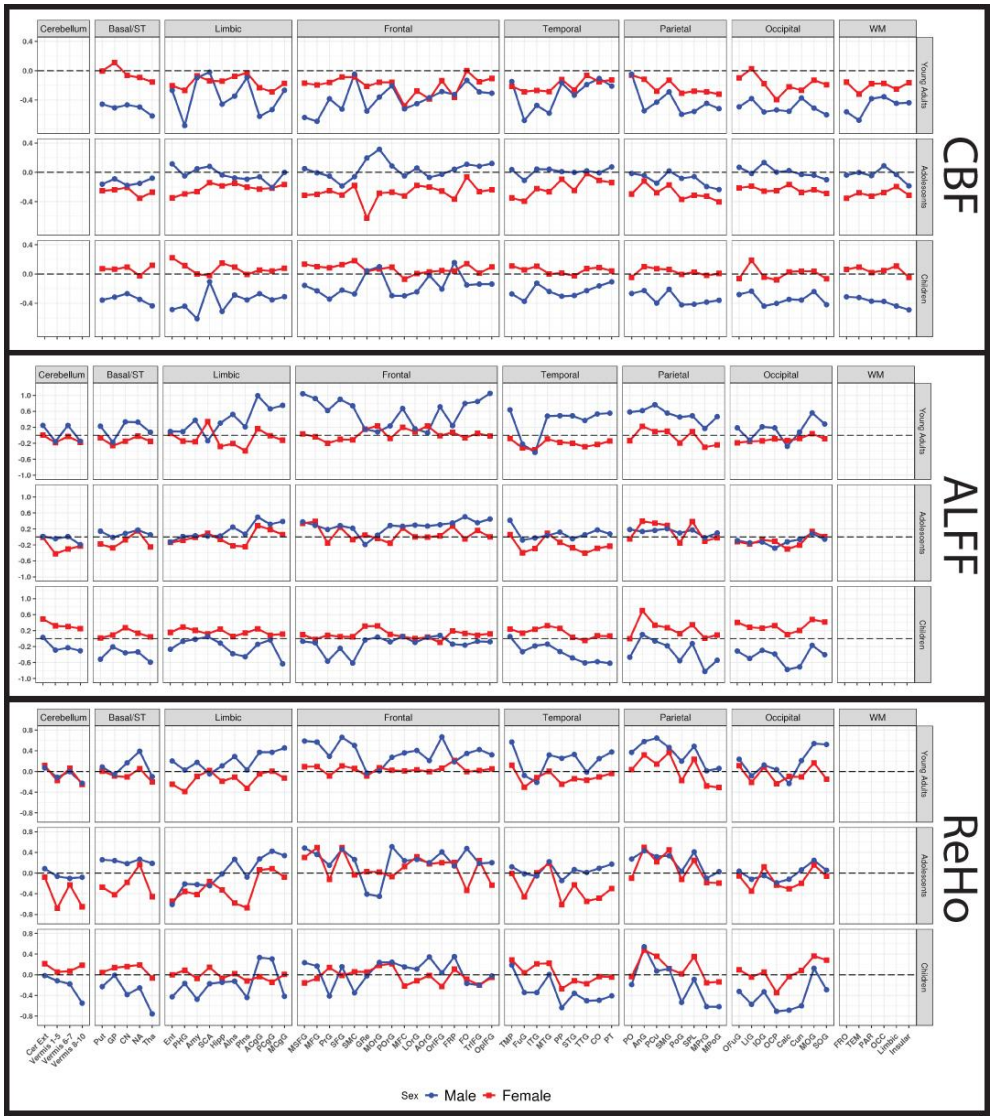


Figure 4. Effect sizes (Cohen D) for the physiologic parameters CBF (top panel), ALFF (middle panel) and ReHo (bottom panel) for each age group.

For ALFF, there was a significant performance by brain section interaction that also differed by age and sex (GEE, Wald $\chi2 = 66.42$, df=35, p =0.0011) justifying exploratory analyses within each brain section. The exploratory analyses revealed significant four way interactions of Performance*Region*Age*Sex, in all but cerebellum (Table S3), which did show a marginal Performance*Region interaction (p=0.0314). These interactions are illustrated in Figure 4, middle panel. For males, higher ALFF in the high performing groups develops from childhood through adolescence to adulthood. Indeed, high performing children have lower ALFF

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

than their low performing counterparts, but the adult groups are characterized by robust effect sizes, reaching and exceeding .8 SDs in some frontal (medial superior frontal and opercular inferior frontal gyrus) and parietal (precuneus) regions. In females, by contrast, effect sizes are considerably smaller (+/- .4 SDs) and mature in the opposite direction to that of males, with adult high performers having *lower* ALFF in most regions.

For ReHo, there was a significant performance by brain section interaction (GEE, Wald $\chi 2 = 73.41$, df=35, p =0.0002). This analysis also yielded significant interactions with sex and age. The exploratory analyses within a brain section revealed significant 4 way interactions of Performance*Region*Age*Sex, in all brain sections except occipital (Table S3). The occipital regions did show a significant Performance by Region interaction. These interactions are illustrated in Figure 4, bottom panel. As with ALFF, in males the effect sizes became more positive from childhood to adulthood and exceed .7 SDs in frontal (superior frontal and orbital inferior frontal gyrus) and a parietal (precuneus) region. In females, as for ALFF, effect sizes were smaller and became more negative in adults compared to children in most regions.

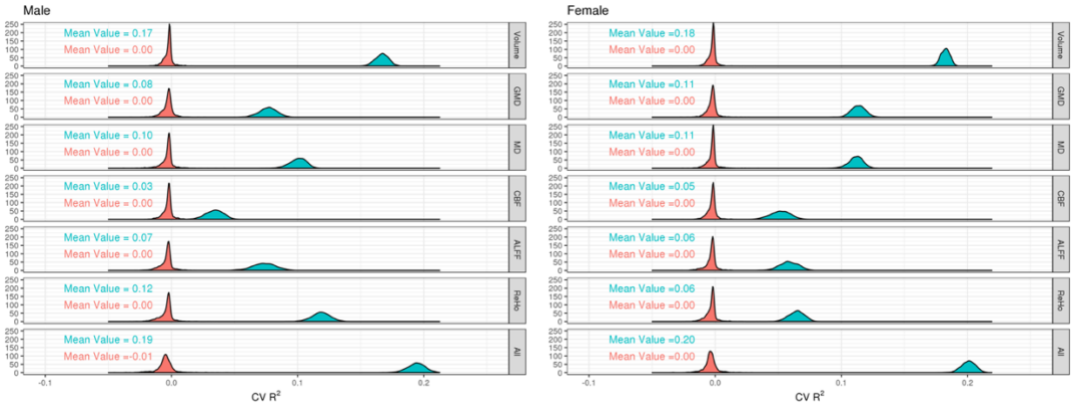
Data-driven analyses

In the data-driven analysis, we applied ridge regression (Hoerl & Kennard, 1970) to the multimodal brain parameters with all regions allowed to compete for explaining variance in the cognitive performance measure. We employed ridge regression in a cross-validated manner. Data were randomly split into training (90%) and testing (10%) sets; models were built in the training set, which includes optimizing the penalty as well as determining the model coefficients; and this trained model was applied to the testing set, resulting in an out-of-sample measure of percent variance explained (R^2). The above steps were repeated 1000 times, and the R^2 values reported below are the means of these 1000 splits. The same procedure was repeated on the same

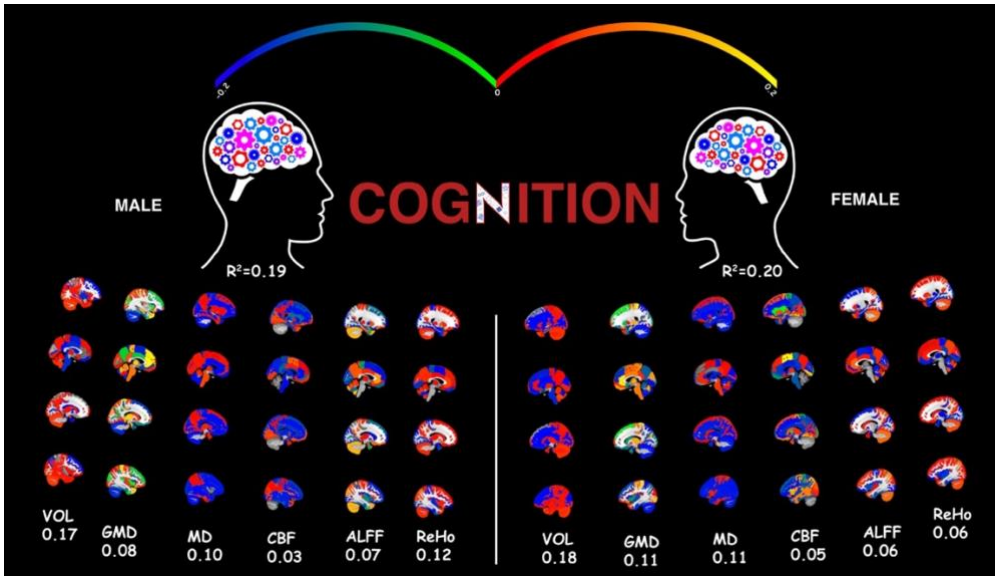
bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license

data, randomly permuting the performance parameter among subjects to generate a null distribution. Notably, all real distributions were significantly greater than the null distributions. As can be seen in Figure 5, volume alone explained the largest portion of variance in performance, producing a cross-validated R^2 that exceeds 17% of variance in both males and females. Other parameters explained much smaller portions of the variance, with GMD, MD and CBF explaining similar proportions of variance in males and females, in the range of 4 to 12%, while ALFF and ReHo explained larger proportion of variance in males than in females. Notably, when all modalities were entered, the ridge regression produced cross-validated R^2 of 20% for both males and females.

a.



b.

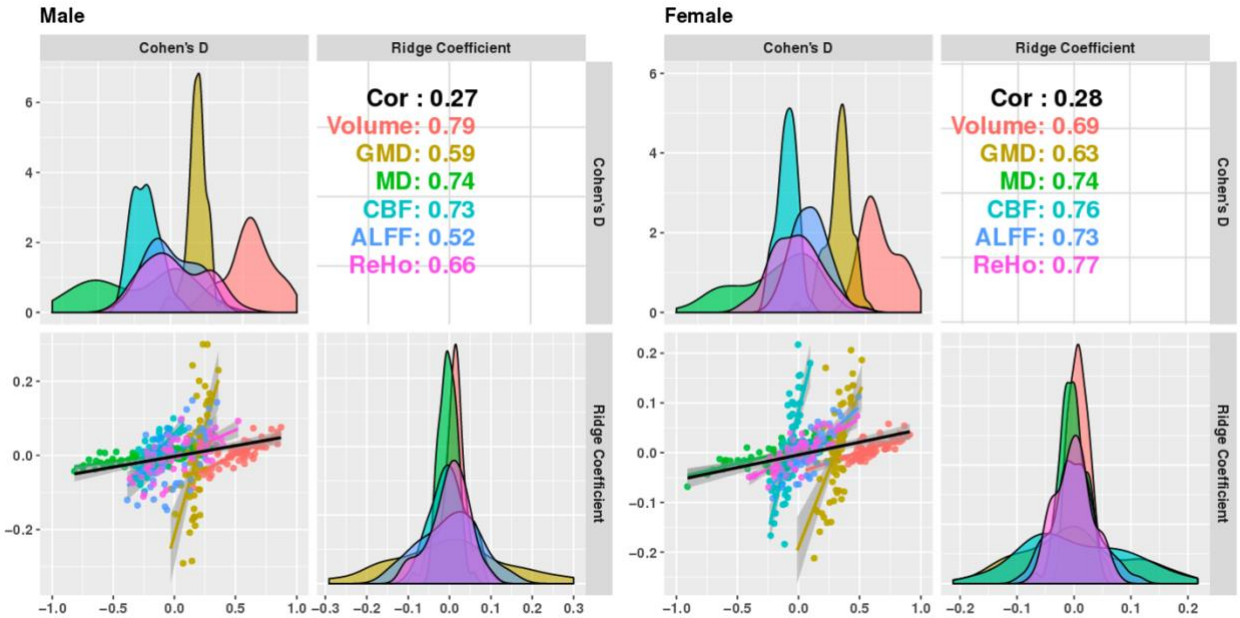


bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5. a. Results of ridge regressions predicting performance for each modality and across modalities. The distribution are shown of cross-validated R^2s for 1000 executions on real (blue bars) and permuted (red bars) data. b. Images showing regional correlations with performance for each parameter by sex.

Integration of hypothesis and data-driven results

To examine the relation between the effect sizes produced by contrasting the high and low performance groups and the results of the data-driven ridge regressions, we correlated the effect sizes with the ridge regression coefficients. Ridge coefficients were obtained by training a model in the entire dataset. The results (Figure 6) indicate that the overall correlation was moderate (0.27 in males and 0.28 in females), but this masked high within-modality correlations (ranging from .55 to .79 in males and from .62 to .79 in females). The attenuated overall correlations seems to reflect an amalgamation "Yule-Simpson" effect (Yule 1903, Blyth 1972), where the overall lower correlation masks underlying high correlations within each brain parameter. Thus, within each parameter (volume, GMD, etc.), the regions selected as most important by the ridge regression also had larger effect sizes when contrasting high and low performance groups.



bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 6. Scatterplot of effect sizes contrasting High and Low Performance groups with "Importance Parameters" from Ridge Regressions predicting performance for each modality and across modalities.

DISCUSSION

Our results offer fresh insights regarding how cognition is related to multimodal parameters of brain structure and function. Volume was far and away the brain parameter most strongly associated with cognitive performance, confirming earlier findings associating higher brain volumes with better cognitive abilities (Witelson, Beresh, & Kigar, 2005; Gignac and Bates, 2017; Nave et al., 2018; Pietschnig et al., 2015). In our sample, effect sizes separating high from low performance groups were moderate to large, and cross-validated R^2s for predicting performance based on volume alone exceeded .17 for both males and females. This estimate of explained variance is at the middle-range of estimates from prior studies, which vary from 3% to >30%. We also established that high GMD relates to better performance, although the effect sizes and variance explained were more modest (around 7% for males and 12% for females). Diffusion-weighted parameters showed that high performance was associated with lower MD values. While these effects were regionally specific, they reached moderate to large effect sizes (.7 to .8 SDs) in some frontal and temporal regions, and overall explained over 10% of performance variance in both males and females. Thus, brains of high performing individuals are anatomically characterized by larger volumes and denser and less coarsely organized gray matter. The neuroanatomic performance associations are stable across the current age range and in both males and females for volume, or become more pronounced across developmental epochs of childhood, adolescence and young adulthood for GMD and MD.

We also uncovered how physiologic parameters relate to performance. Resting CBF values were negatively associated with performance, showing small effect sizes (.2 to .4 SDs) and explaining about 3% and 5% of performance variance in males and females, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Notably, although effect sizes of lower CBF in high performers reached maturity levels earlier in females than in males (by adolescence), they ended up considerably larger in the young adult males. While we had no prior studies on which to base a hypothesis on whether high performance in this developmental sample should be associated with higher or lower CBF, it is notable that CBF was measured at a resting state, characterized as the "default mode" (Raichle et al., 2001) condition. Our finding that lower resting-state CBF is associated with better overall performance is consistent with reports that deactivation of the default-mode network during task performance is as predictive of performance as activation of task-related regions (Satterthwaite et al., 2013). A lower "idling rate" may be conducive to better performance by permitting activation when the individual is faced with a task while preserving energy in the absence of a task.

ALFF and ReHo derived from resting state fMRI, showed more complex associations with performance, which differed in males and females. Males showed higher values associated with good performance, a relationship that strengthened from childhood to adolescence to early adulthood. In contrast, in females these effects were smaller, and in adults, high performers had even lower values in most regions. It is unclear why higher or lower amplitude of low frequency fluctuations or regional homogeneity values would relate to cognitive performance, and our findings should motivate further study with complementary methodologies and analytic approaches.

The results overall indicate that optimal cognitive performance is supported by fine-tuning of regionally distributed anatomic and physiologic parameters. Notably, these relationships between anatomic brain parameters and performance are generally stable within this wide developmental age range of 8 to 22 years, notwithstanding the major effects of age on

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

all parameters involved. For the anatomic parameters, the relation with performance is stable or becomes more pronounced with maturation. The physiologic parameters are also associated with performance, but these associations evolve during this age span and differ for males and females.

The regional distribution of performance group differences varied by modality, but across modalities it was more pronounced for some regions. The regions that were most strongly modulated in relation to performance included few subcortical regions, specifically thalamus, hippocampus, parahippocampal gyrus and entorhinal cortex. Subcortical regions have been implicated in higher-order cognitive function and memory (Koziol et al., 2014; Münte et el., 2008; Wolff & Vann, 2019). Effect sizes were generally strongest for cortical GM regions, and these included frontal (orbital and midfrontal) temporal (midtemporal), Parietal (precuneus and superior parietal) and occipital cortex (fusiform and lingual). All these regions have been implicated as important nodes in complex cognition (Gur et al., 1999; Dubois, Galdi, Paul, Adolphs, 2018; Nave, Jung, Karlsson, Kable, Koellinger, 2019).

Although the main findings were robust across age groups, sex differences modulated observed effects in informative ways. The sex differences overall indicated complementary mechanisms in males and females that serve to compensate for sex differences in brain parameters. Sex differences in brain-behavior associations have been well documented (e.g., Faraone and Tsuang, 2001, Gur et al., 1982, 1999, 2000; Jazin and Cahill, 2010, Ragland et al., 2000, Raznahan et al., 2011, Satterthwaite et al., 2015). Thus, lower volume in females is compensated for by higher GMD and performance is further modulated by MD, CBF, ALFF and ReHo, which may together account for equal cognitive performance. Furthermore, age-related differences in this developmental cohort were smaller for females than for males, indicating that females reach adult differences earlier. Indeed, when we applied age prediction models based on

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

volume and MD, chronological age was over-predicted by brain age in the younger age range (8 to 16 years) in females compared to males, and this effect reversed toward young adulthood, where females had "younger looking" brains (Erus et al., 2014). This sex difference in adult samples seems to extend across metabolic parameters and the age range of 20 to 80 years - females consistently show younger brain age than male counterparts with the same chronological age (Goyal et al., 2019). We might speculate that earlier stabilization of metabolic parameters in females helps sustain brain integrity throughout the adult lifespan. Such a complementarity between the sexes might have enhanced survival and reproduction in humans' environment of evolutionary adaptedness (Barkow, Cosmides, & Tooby, 1995). These sex differences could further reflect complementary reliance on different aspects of brain structure and function to optimize cognitive performance.

The data-driven analyses indicated that brain parameters of structure and function can predict a fairly high portion of the variance in cognitive performance. While the strongest predictor was volume, which alone accounted for about 17% of the variance, even the parameter with lowest predictive value accounted for ~5% of the variance. Using a combined set of parameters we were able to explain nearly 20% of the variance (cross-validated) in both males and females. Most previous studies to which we can compare our results examined volume. Our R^2 values are considerably higher than those reported for volumes by Nave et al., (2018), which estimate that volume explains slightly over 3% of the variance in cognitive performance. Possibly the more extensive battery on which our performance measure was based, as well as the use of the same scanner, could have eliminated some sources of noise in estimating the dependent measures. Overall, considering the inherent error in all our measurements, these results indicate substantial coupling between cognitive performance and parameters of brain

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

structure and function. Additional variance can be explained by more refined cognitive parameters that relate to specific brain systems, by examination of brain networks, or by measures such as regional activation concurrent with performance (Roalf et al., 2014). Indeed, earlier work has shown that task-activated fMRI is a better predictor of performance than the resting-state measures examined here (Greene, Gao, Scheinost, & Constable, 2018; Yoo et al., 2018).

Several limitations of this study are noteworthy. The age range of the sample, 8 to 22 years, limits generalizability to other ages. Within this age range, in which changes occur in all brain modalities examined, age did not generally affect the performance-related differences. Another limitation of the study was the focus on a single parameter of cognitive capacity. This focus was necessitated by the complexity of probing multiple brain regions across modalities and accounting for age effects and sex differences. The measure selected is the closest proxy for "IQ", which was used in other studies and thus improves comparability of results. Future analyses can focus on other performance domains, such as episodic memory and social cognition, and more specific aspects of performance such as accuracy compared to speed. The study is also limited by analyzing data across the entire sample, which is quite heterogeneous and, while ascertained through general pediatric services and not psychiatric services, still included individuals with significant psychopathology (Calkins et al., 2015) and adverse life events (Barzilay et al., 2018), and from diverse sociodemographic, ethnic and racial backgrounds. Perhaps stronger and more coherent effects could be seen if we limited the analyses to the subsample of typically developing youth without any significant disorder or to more homogeneous populations. We believe that while such analyses have merit and could reveal effects of different disorders on the observed relationships, the heterogeneity and diversity

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of our sample enhance generalizability of the reported results. Finally, our analyses examined regional parameters of brain structure and function and related them to individual differences in behavioral measures of cognitive performance taken within the same timeframe, but not contemporaneously in the scanner. More variance in behavioral measures could be linked to functional brain parameters acquired contemporaneously (Roalf et al., 2014).

Notwithstanding its limitations, the present study provides some "benchmarks" for assessing relations among brain parameters and performance. The results can guide hypotheses on how brain structure and function relate to individual differences in cognitive capacity, and offer the ability to gauge the relevance to cognitive performance of group differences or changes in brain parameters. Furthermore, acquisition of each parameter is costly in time and data management resources, and our study can inform design of future large-scale neuroimaging studies based on the relevance of associating acquired brain parameters with cognitive performance. Our finding of lower resting CBF in high performers is worthy of special emphasis since, unlike the structural parameters of volume, GMD and MD, it relates to brain function. Uncovering a physiologic index associated with individual differences in cognitive performance has important implications for developing a scientific basis for social and medical prevention, education and intervention strategies. Anatomy is unlikely to be readily affected by behavioral or pharmacologic treatment. By contrast, physiologic states such as measured by CBF, ALFF and ReHo, can be changed within seconds, and it is easier to conceive of treatments that can affect resting-state CBF for sustainable durations. Notably, current methods for rehabilitation of brain dysfunction emphasize activation of task-related brain systems. Our findings that lower resting CBF and increased resting state connectivity are associated with better performance suggest that emphasis should also be placed on training in deactivation of task-relevant regions in the absence

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of a task. Indeed, our results may offer a scientific basis for the benefits of procedures such as meditation, which emphasize relaxation associated with the absence of goal-oriented behavior and results in reduced default-mode activity (Brewer et al., 2011; Hasenkamp and Barsalou, 2012). That both greater volume and gray matter density of brain and lower basal metabolic rate are associated with cognitive abilities is consistent with preservation of tissue at low energy consumption as the "holy grail" for optimal organ function.

ACKNOWLEDGMENTS

We thank the participants of the Philadelphia Neurodevelopmental Cohort and the members of the Recruitment, Assessment, Neuroimaging and Data Teams whose contributions made this project possible. This work was supported by NIH grant MH107235, MH089983, MH096891, MHP50MH06891, R01MH113550, R01MH112847, the Dowshen Neuroscience fund, and the Lifespan Brain Institute of Children's Hospital of Philadelphia and Penn Medicine, University of Pennsylvania.

AUTHOR CONTRIBUTIONS

R.E.G. and R.C.G. conceived the project, designed the study, guided data analysis, interpreted the results and wrote the manuscript. M.A.E., R.V. and J.A.D. designed the imaging protocol and participated in data analysis, R.C.G., T.M.M., A.F.G.R., A.P. and K.R. analyzed data, made figures and wrote sections of the manuscript. T.D.S., D.R.R., D.H.W., R.V., C.D. and E.D.G. guided image processing and interpretation of results. W.B.B. and R.T.S. guided statistical analysis. All authors reviewed and contributed to the write-up of the manuscript.

DECLARATION OF INTERESTS

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license

All authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. 2011a. A reproducible evaluation of ANTs similarity metric performance in brain image registration. NeuroImage. 54: 2033–2044.

Avants BB, Tustiso NJ, Wu J, Cook PA, Gee JC. 2011b. An open source multivariate framework for n-tissue segmentation with evaluation on public data. Neuroinformatics. 9: 381–400.

Barkow JH, Cosmides L, Tooby J. (Eds.). 1995. The adapted mind: Evolutionary psychology and the generation of culture. Oxford University Press, USA.

Barzilay R, Calkins ME, Moore TM, Wolf DH, Satterthwaite TD, Scott CJ, Jones JD, Benton TD, Gur RC, Gur RE. 2019. Association between traumatic stress load, psychopathology, and cognition in the Philadelphia Neurodevelopmental Cohort. Psychol Med. 49:325-334.

Blyth CR. 1972. On Simpson's Paradox and the Sure-Thing Principle. Journal of the American Statistical Association. 67: 364–366.

Brewer JA, Worhunsky PD, Gray JR, Tang YY, Weber J, Kober H. 2011. Meditation experience is associated with differences in default mode network activity and connectivity. Proc. Natl. Acad. Sci. U. S. A. 108: 20254-20259.

Calkins ME, Merikangas KR, Moore TM, Burstein M, Behr A., Satterthwaite TD, Ruparel K, Wolf DH, Roalf DR, Mentch FD, et al. 2015. The Philadelphia Neurodevelopmental Cohort: constructing a deep phenotyping collaborative. J. Child Psychol. Psychiatr 56: 356-1369.

Ciric R., Wolf DH, Power D. Roalf, DR, Baum GL, Ruparel K, Shinohara RT, Elliott MA, Eickhoff S, Davatzikos C, Gur RC, Gur RE, Bassett DS, Satterthwaite TD. 2017. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. NeuroImage. 154: 174-187.

Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S. 1999. Sources of mathematical

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

thinking: behavioral and brain-imaging evidence. Science. 284: 970-974.

Dubois J, Galdi P, Paul LK, Adolphs R. 2018. A distributed brain network predicts general intelligence from resting-state human neuroimaging data. Philos. Trans. R. Soc. Lond. B Biol. Sci. 373:(1756).

Forster MR. 2000. Key concepts in model selection: performance and generalizability. J. Math. Psychol. 44: 205-231.

Friedman J, Hastie T, Simon N, Qian J, Tibshirani R. 2017. glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models.

Erus G, Battapady H, Satterthwaite TD, Hakonarson H, Gur RE, Davatzikos C, Gur RC. 2014. Imaging patterns of brain development and their relationship to cognition. Cerebral Cortex. 25: 1676-1684.

Finn E., Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable, RT. 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat Neurosci. 18: 1664-1671.

Fong AHC, Yoo K, Rosenberg MD, Zhang S, Li CR, Scheinost D, Constable RT, Chun MM. 2019. Dynamic functional connectivity during task performance and rest predicts individual differences in attention across studies. Neuroimage. 188:14-25.

Genç E, Fraenz C, Schlüter C, Friedrich P, Hossiep R, Voelkle M, Ling, JM, Güntürkün O, Jung, RE. 2018 Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence. Nat Commun. 9: 1905 (1-12).

Gennatas ED, Avants BB, Wolf DH, Satterthwaite TD, Ruparel K, Ciric R, Hakonarson H, Gur RE, Gur RC. 2017. Age-related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood. J. Neurosci. 37: 5065-5073.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gignac GE, Bates TC.2017). Brain volume and intelligence: The moderating role of intelligence measurement quality. Intelligence. 64:18–29.

Goldstein JM, Seidman LJ, Horton NJ, Makris N, Kennedy DN, Caviness VS Jr, Faraone SV, Tsuang MT. 2001. Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. Cereb. Cortex. 11: 490-497.

Goyal MS, Blazey TM, Su Y, Couture LE, Durbin TJ., Bateman RJ, Benzinge, TL, Morris JC, Raichle ME, Vlassenko AG. 2019. Persistent metabolic youth in the aging female brain. Proc Natl Acad Sci U S A. 116: 3251-3255.

Greene AS, Gao S, Scheinost D, Constable RT. 2018. Task-induced brain state manipulation improves prediction of individual traits. Nature comm. 9: 2807.

Greve DN, Fischl B. 2009. Accurate and robust brain image alignment using boundary-based registration. NeuroImage. 48: 63–72.

Gu R.C, Gur RE, Obrist WD, Hungerbuhler JP, Younkin D, Rosen AD, Skolnick BE, Reivich M. 1982. Sex and handedness differences in cerebral blood flow during rest and cognitive activity. Science. 217: 659-661.

Gur RC, Turetsky BI, Matsui M, Yan M, Bilker W, Hughett P, Gur RE. 1999. Sex differences in brain gray and white matter in healthy young adults: correlations with cognitive performance. J. Neurosci. 19: 4065-4072.

Gur RC, Alsop D, Glahn D, Petty R, Swanson CL, Maldjian JA, Turetsky BI, Detre JA, Gee J, Gur RE. 2000. An fMRI study of sex differences in regional activation to a verbal and a spatial task. Brain Lang. 74: 157-170.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB, Brensinger C, Gur RE. 2010. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. J Neurosci Methods. 187: 254-262.

Gur C ,Richard J, Calkins ME, Chiavacci R, Hansen JA, Bilker B., Loughead , Connolly JJ, Qiu H, Mentch FD, et al. 2012. Age group and sex differences in performance on a computerized neurocognitive battery in children age 8-21. Neuropsychol. 26: 251-265.

Gur RC, Calkins ME, Satterthwaite TD, Ruparel K, Bilker WB, Moore TM, Savitt AP, Hakonarson H, Gur RE. 2014. Neurocognitive growth charting in psychosis spectrum youths. JAMA Psychiatry. 71: 366-374.

Hasenkamp W, Barsalou LW. 2012. Effects of meditation experience on functional connectivity of distributed brain networks. Front. Hum. Neurosci. 1: 38.

Hoerl AE, Kennard RW. 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics. 12: 55-67.

Hshieh TT, Dai W, Cavallari M, Guttmann CR, Meier DS, Schmitt EM, Dickerson C., Pres DZ, Marcantonio ER, Jones R., et al. 2017. Cerebral blood flow MRI in the nondemented elderly is not predictive of post-operative delirium but is correlated with cognitive performance. J. Cereb. Blood Flow Metab. 37: 1386-1397.

Ingalhalikar M, Smit A., Parker D, Satterthwaite TD, Elliott MA, Ruparel K, Hakonarson H, Gur RE, Gur RC, Verma R. 2014. Sex differences in the structural connectome of the human brain. Proc. Natl. Acad. Sci. U. S. A. 111: 823-828.

Jain V, Duda J, Avants B, Giannetta M, Xie X, Roberts T, Detre JA, Hurt H, Wehrli FW, Wang DJJ. 2012. Longitudinal reproducibility and accuracy of pseudo-continuous arterial spin-labeled perfusion MR imaging in typically developing children. Radiology. 263, 527–536.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Jangraw DC, Gonzalez-Castillo J, Handwerker DA, Ghane M, Rosenberg MD, Panwar P, Bandettini PA. 2018. A functional connectivity-based neuromarker of sustained attention generalizes to predict recall in a reading task. Neuroimage. 166: 99-109.

Jazin E, Cahill L. 2010. Sex differences in molecular neuroscience: from fruit flies to humans. Nat. Rev. Neurosci. 11: 9-17.

Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage. 17: 825–841.

Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. 2012. FSL. NeuroImage. 62: 782–790.

Klein A, Ghos SS, Avants B, Yeo BTT, Fischl B, Ardekani B, Gee JC, Mann JJ, Parsey RV. 2010. Evaluation of volume-based and surface-based brain image registration methods. NeuroImage. 51: 214–220.

Koziol LF, Budding D, Andreasen N, D'Arrigo , Bulgheroni, S, Imamizu H, Ito M, Manto M, Marvel C, Parker K, et al. 2014. Consensus paper: the cerebellum's role in movement and cognition. Cerebellum. 13: 151-177.

Kuhn M, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, Team, the RC, Benesty M, Lescarbeau R, Ziem A, Scrucca L, Tang Y, Candan C, Hunt T. 2016. caret: Classification and Regression Training.

Marcus D, Wang, TH, Parker J, Csernansky JG, Morris JC, Buckner RL. 2007. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J. Cogn. Neurosci. 19: 1498–1507.

Moore TM, Reise SP, Gur RE, Hakonarson H, Gur RC. 2015. Psychometric properties of the Penn Computerized Neurocognitive Battery. Neuropsychol. 29: 235-246.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Münte TF, Heldmann M, Hinrichs H, Marco-Pallares J, Krämer UM, Sturm V, Heinze H J. 2008. Contribution of subcortical structures to cognition assessed with invasive electrophysiology in humans. Frontiers in neuroscience. 2: 72–78.

Nave G, Jung WH, Karlsson Linnér R, Kable JW, Koellinger PD. 2019. Are bigger brains smarter? Evidence from a large-scale preregistered study. Psychol. Sci. 30: 43-54.

Pamplona GS, Santos Neto GS, Rosset SR, Rogers BP, Salmon CE. 2015. Analyzing the association between functional connectivity of the brain and intellectual performance. Front Hum Neurosci. 9: 61.

Pietschnig J, Penke L, Wicherts JM, Zeiler M, Voracek M. 2015. Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean? Neurosci. Biobehav. Rev. 57: 411-432.

Qiu D, Tan L-H, Zhou K, Khong PL. 2008. Diffusion tensor imaging of normal white matter maturation from late childhood to young adulthood: Voxel-wise evaluation of mean diffusivity, fractional anisotropy, radial and axial diffusivities, and correlation with reading development. NeuroImage. 41: 223–232.

R Core Team (2015). R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. 2001. A default mode of brain function. Proc. Natl. Acad. Sci. U.S.A. 98: 676-682.

Ragland JD, Coleman AR, Gur RC, Glahn DC, Gur RE. 2000. Sex differences in brain-behavior relationships between verbal episodic memory and resting regional cerebral blood flow. Neuropsychol. 38: 451-461.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rane S, Koh N, Boord P, Madhyastha T, Askren MK, Jayadev S, Cholerton B, Larson E, Grabowski TJ. 2018. Quantitative cerebrovascular pathology in a community-based cohort of older adults. Neurobiol. Aging. 65: 77-85.

Raznahan A, Shaw P, Lalonde F, Stockman M, Wallace GL, Greenstein D, Clasen L, Gogtay N, Giedd JN. 2011. How does your cortex grow? J. Neurosci. 31: 7174-7177.

Ritchie SJ, Booth T, Valdés Hernández MD, Corley J, Maniega S., Gow AJ, Royle NA, Pattie, A, Karama S, Starr JM, et al. 2015. Beyond a bigger brain: Multivariable structural brain imaging and intelligence. Intelligence. 51: 47-56.

Roalf DR, Gur RE, Ruparel K, Calkins M., Satterthwaite TD, Bilker WB, Hakonarson H, Harris LJ, Gur RC. 2014. Within-individual variability in neurocognitive performance: age- and sex-related differences in children and youths from ages 8 to 21. Neuropsychol. 28: 506-518.

Roalf DR, Ruparel K, Gur RE, Bilker W, Gerraty R, Elliott MA, Gallagher RS, Almasy L, Pogue-Geile MF, Prasad K, et al. 2014. Neuroimaging predictors of cognitive performance across a standardized neurocognitive battery. Neuropsychol. 28: 161-176.

Ryman SG, Yeo RA, Witkiewitz K, Vakhtin AA, van den Heuvel M, de Reus M, Flores RA, Wertz CR, Jung RE. 2016. Fronto-Parietal gray matter and white matter efficiency differentially predict intelligence in males and females. Hum Brain Mapp. 37: 4006-4016.

Roalf DR, Quarmley M, Elliott MA, Satterthwaite TD, Vandekar SN, Ruparel K, Gennatas ED, Calkins ME, Moore TM, Hopson R, Prabhakaran K, Jackson CT, Verma R, Hakonarson H, Gur RC, Gur RE. 2016. The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. NeuroImage. 125: 903–919.

Rosen A, Roalf DR, Ruparel K, Blake J, Seelaus K, Villa LP, Ciric R, Cook PA, Davatzikos C, Elliott MA, Garcia De La Garza A, Gennatas ED, Quarmley M, Schmitt JE, Shionhara RT,

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tisdall MD, Craddock RC, Gur RE, Gur RC, Satterthwaite TD. 2017. Data-driven assessment of structural image quality. bioRxiv.

Satterthwaite TD, Elliott MA, Ruparel K, Loughead J, Prabhakaran K, Calkins ME, Hopson R, Jackson C, Keefe J, Riley M, Mentch FD, Sleiman P, Verma R, Davatzikos C, Hakonarson H, Gur RC, Gur RE. 2014. Neuroimaging of the Philadelphia neurodevelopmental cohort. Neuroimage. 86:544-553.

Satterthwaite TD, Wolf DH, Erus G, Ruparel K, Elliott MA, Gennatas ED, Hopson, R, Jackson C, Prabhakaran K, Bilker WB, et al. 2013. Functional maturation of the executive system during adolescence. J. Neurosci. 33: 16249-16261.

Satterthwaite TD, Elliott MA, Ruparel K, Loughead J, Prabhakaran K, Calkins ME, Hopson R, Jackson C, Keefe J, Riley M, et al. 2014. Neuroimaging of the Philadelphia neurodevelopmental cohort. NeuroImage. 86: 544-553.

Satterthwaite TD, Shinohara RT, Wolf DH, Hopson RD, Elliott MA, Vandekar SN, Ruparel K, Calkins ME, Roalf DR, Gennatas ED, et al. 2014. Impact of puberty on the evolution of cerebral perfusion during adolescence. Proc. Natl. Acad. Sci. U.S.A. 111: 8643-8648.

Satterthwaite TD, Wolf DH, Roalf DR, Ruparel K, Erus G, Vandekar S, Gennatas ED, Elliott MA, Smith A, Hakonarson H, et al. 2015. Linked sex differences in cognition and functional connectivity in youth. Cereb. Cortex. 25: 2383-2894.

Schmithorst VJ, Wilke M, Dardzinski BJ, Holland SK. 2005. Cognitive functions correlate with white matter architecture in a normal pediatric population: a diffusion tensor MR imaging study. Hum. Brain Mapp. 26: 139-147.

Smith SM, Brady JM.1997. SUSAN—A New Approach to Low Level Image Processing. Int. J. Comput. Vis. 23: 45–78.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TE, Glasser MF, Ugurbil K, Barch DM, Van Essen DC, Miller KL. 2015. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat Neurosci. 18: 1565-1567.

Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC., 2010. N4ITK: Improved N3 Bias Correction. IEEE Trans. Med. Imaging. 29:1310–1320.

Tustison NJ, Cook PA, Klein A, Song G, Das SR, Duda JT, Kandel BM, van Strien N, Stone JR, Gee JC, Avants BB. 2014. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. NeuroImage. 99: 166–179.

Vandekar SN, Shinohara RT, Raznahan A, Roalf DR, Ross M, DeLeo N, Ruparel K, Verma R, Wolf DH, Gur RC, et al. 2015. Topologically dissociable patterns of development of the human cerebral cortex. J. Neurosci. 35: 599-609.

Vakhtin AA, Ryman SG, Flores RA, Jung RE. 2014. Functional brain networks contributing to the Parieto-Frontal Integration Theory of Intelligence. Neuroimage. 103: 349-354.

Witelson SF, Beresh H, Kigar DL. 2006. Intelligence and brain size in 100 postmortem brains: sex, lateralization and age factors. Brain. 129: 386–398.

Wang H, Suh JW, Das SR, Pluta JB, Craige C, Yushkevich PA. 2013. Multi-atlas segmentation with joint label fusion. IEEE Trans. Pattern Anal. Mach. Intell. 35: 611–623.

Wang Z, Aguirre GK, Rao H, Wang J, Fernández-Seara MA, Childress AR, Detre JA. 2008. Empirical optimization of ASL data analysis using an ASL data processing toolbox: ASLtbx. Magn. Reson. Imaging. 26: 261–269.

Wolff M, Vann SD. 2019. The cognitive thalamus as a gateway to mental representations. J Neurosci. 39: 3-14.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license

Wu W-., Jain V, Li C, Giannetta M, Hurt H, Wehrli FW, Wang DJJ. 2010. In vivo venous blood T1 measurement using inversion recovery true-FISP in children and adults. Magn. Reson. Med. 64: 1140–1147.

Yoo K, Rosenberg MD, Hsu WT, Zhang S, Li CR, Scheinost D, Constable RT, Chun MM. 2018. Connectome-based predictive modeling of attention: Comparing different functional connectivity features and prediction methods across datasets. Neuroimage. 167: 11-22.

Yule GU. 1903. Notes on the theory of association of attributes in statistics. Biometrika. 2: 121-134.

Zang Y, Jiang T, Lu Y, He Y, Tian L. 2004. Regional homogeneity approach to fMRI data analysis. NeuroImage. 22: 394-400.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

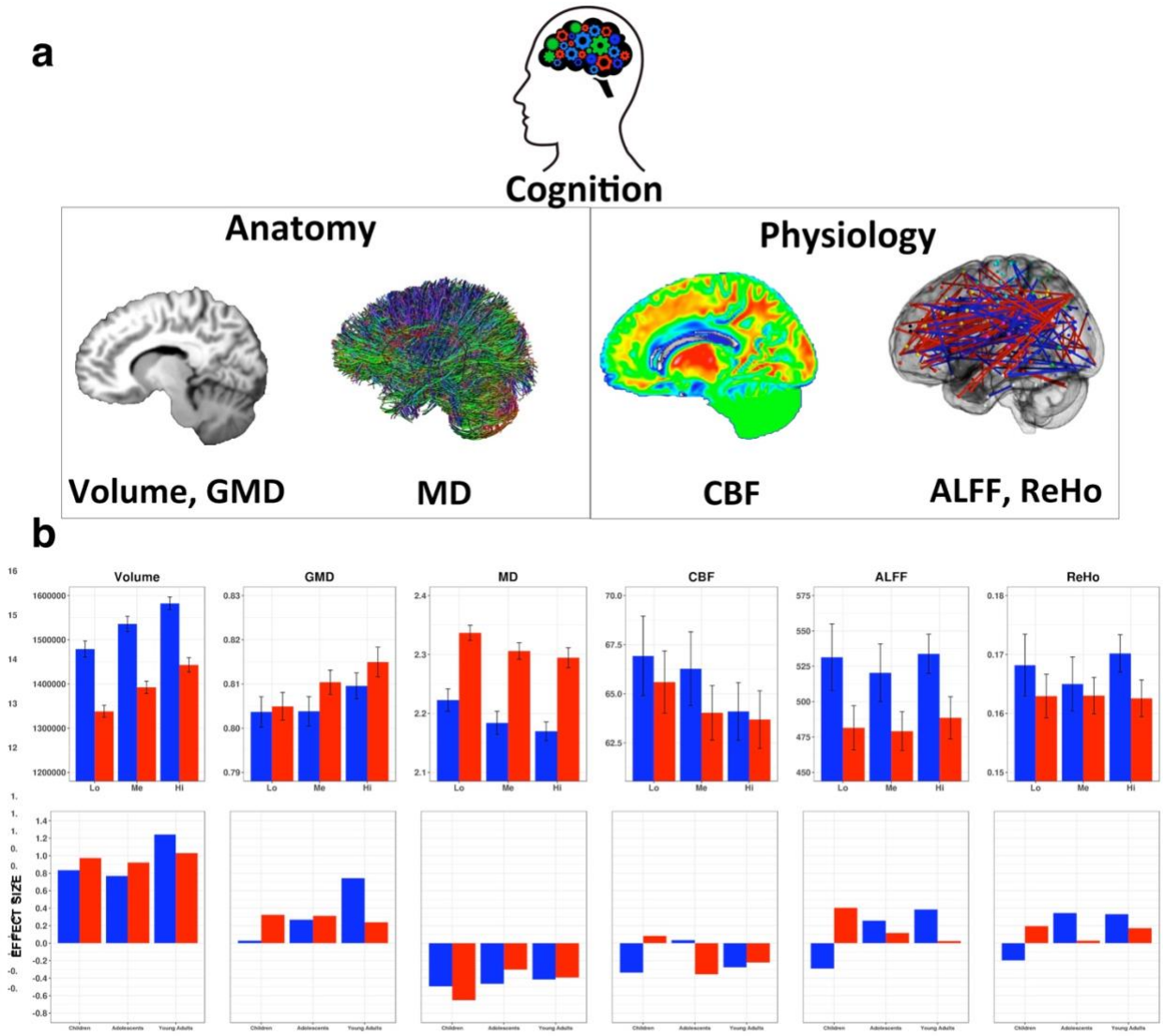


Figure 1. a. Illustration of anatomic and physiologic parameters examined (top panel). b. Global values (±95% CI) for each modality by performance group (tertiles) in males and females. Top panel shows the Performance*Domain interaction displaying means (+/-95% CI) of Low (Lo), Medium (Me) and High (Hi) tertiles of performers on Volume, GMD, MD, CBF, ALFF and ReHo. Bottom panel illustrates the 4-way interaction (Performance*Age*Sex*Domain) by displaying effect sizes (Cohen's D) for the differences between High and Low performers on each parameter.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

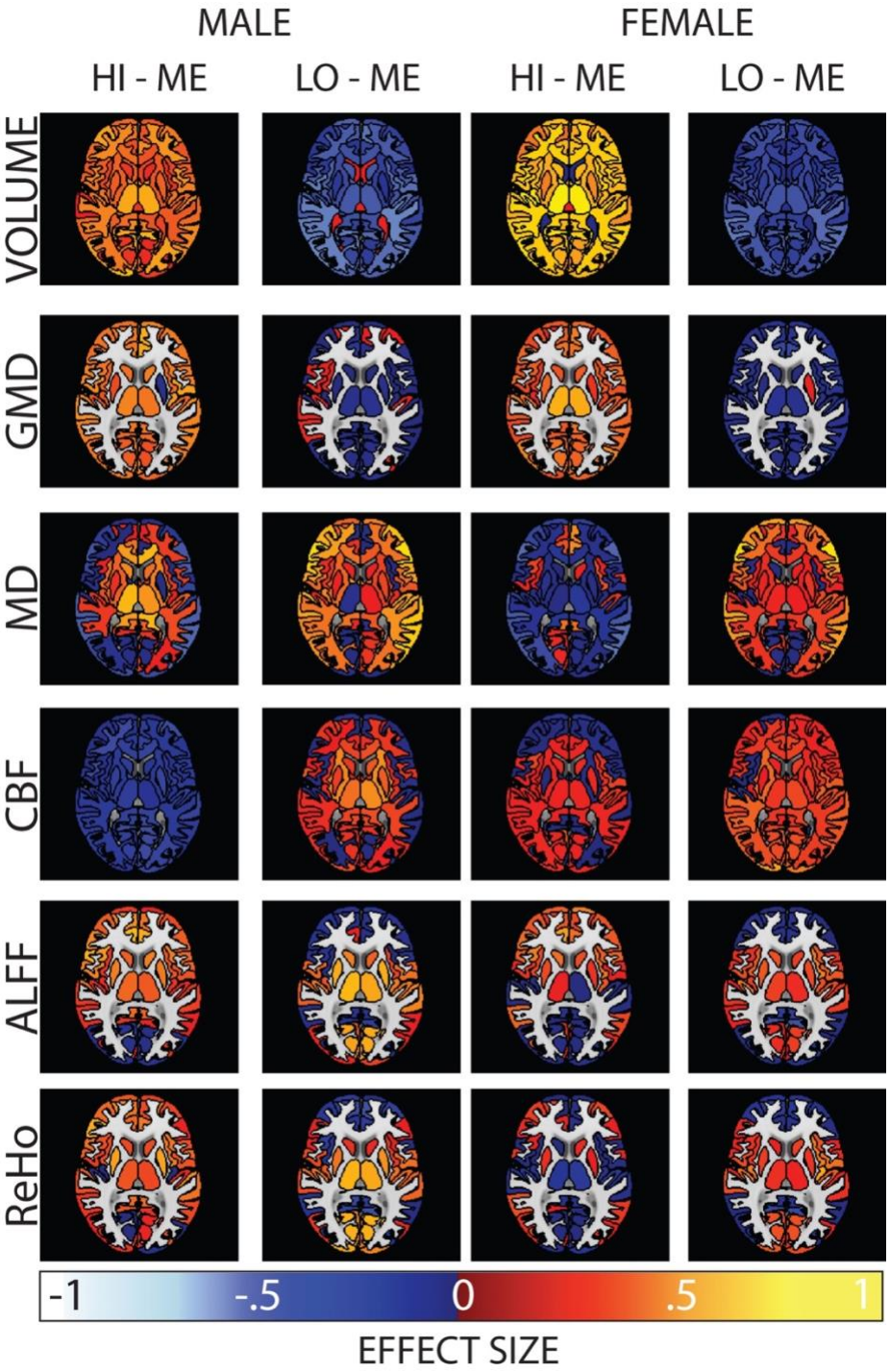


Figure 2. The same axial slice for the six modality contrasts of High vs Medium and Low vs Medium performance group (tertiles) contrast. Effect size units are displayed for high minus medium and low minus medium performance groups in males (left columns) and females (right columns) for volume (top row), GMD (second row), MD (third row), CBF (fourth row), ALFF (fifth row) and ReHo (bottom row). All slices are available as movies here: https://adrose.github.io/hiLo/index.html

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license

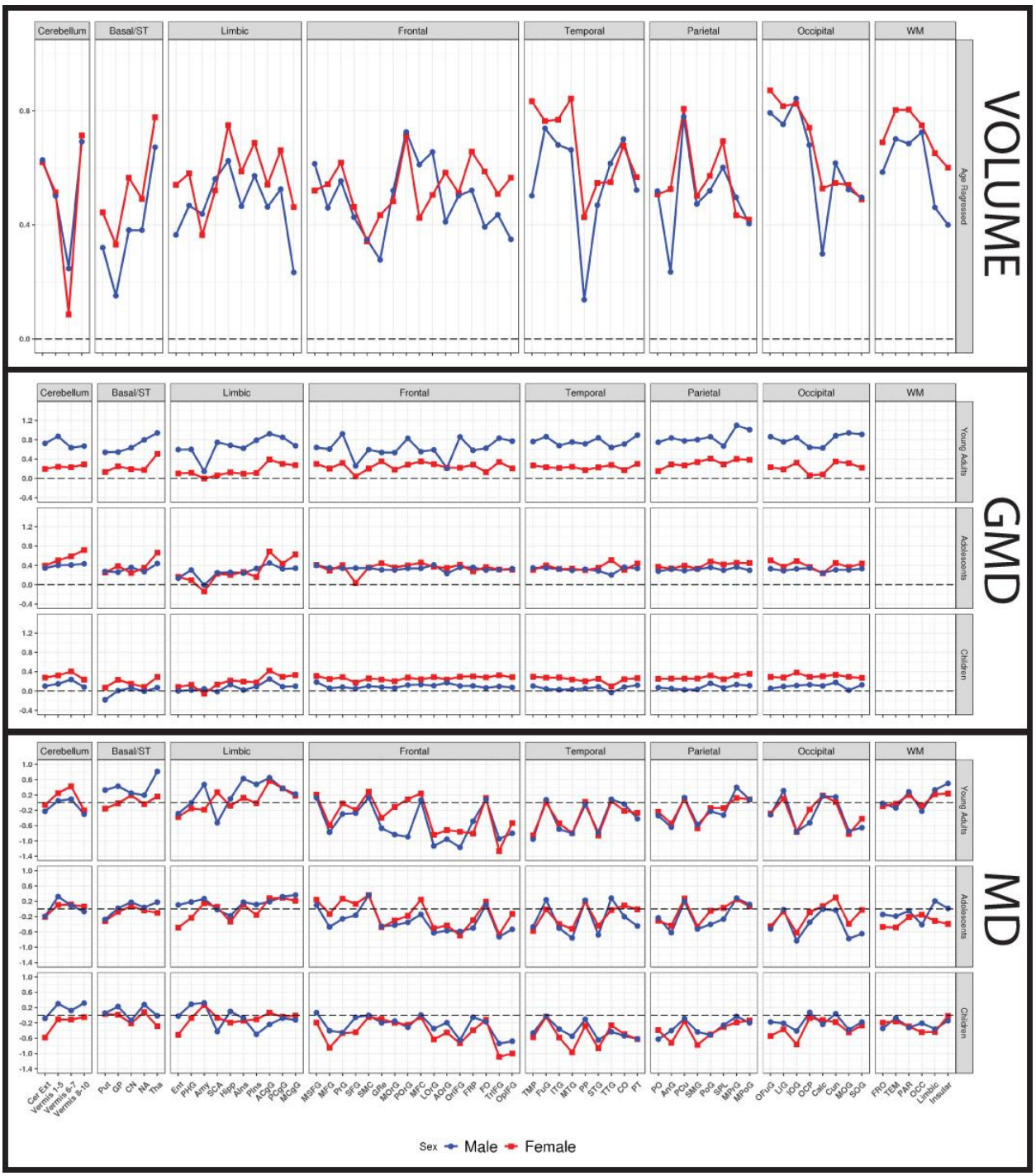


Figure 3. Effect sizes (Cohen D) for Volume, GMD and MD in individual ROIs. Since there was no age interaction for volume, values are shown age-adjusted for the entire sample across the age ranges (top panel). Since there were age interactions for GMD and MD, results are separated by age groups (children, adolescents, young adults) for these parameters (middle and bottom panels). ROI names can be found here: https://pennbbl.github.io/hiLo/oasisLabels/index.html

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

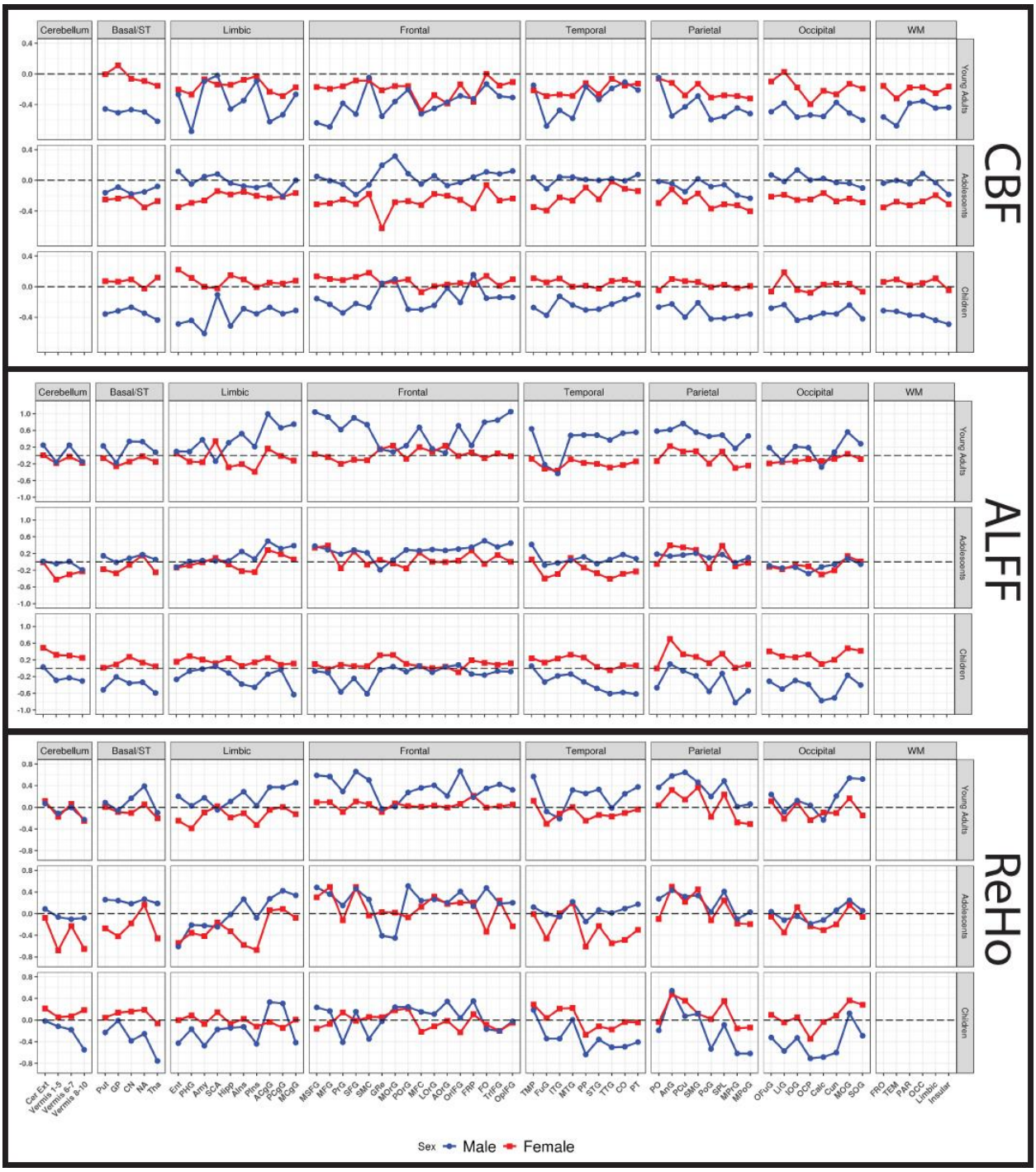
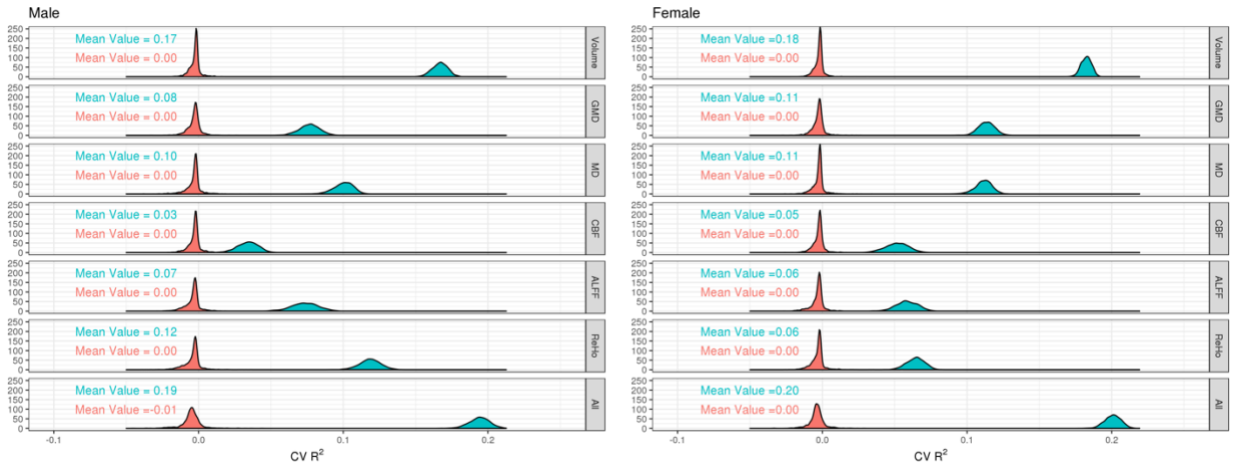


Figure 4. Effect sizes (Cohen D) for the physiologic parameters CBF (top panel), ALFF (middle panel) and ReHo (bottom panel) for each age group.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a.



b.

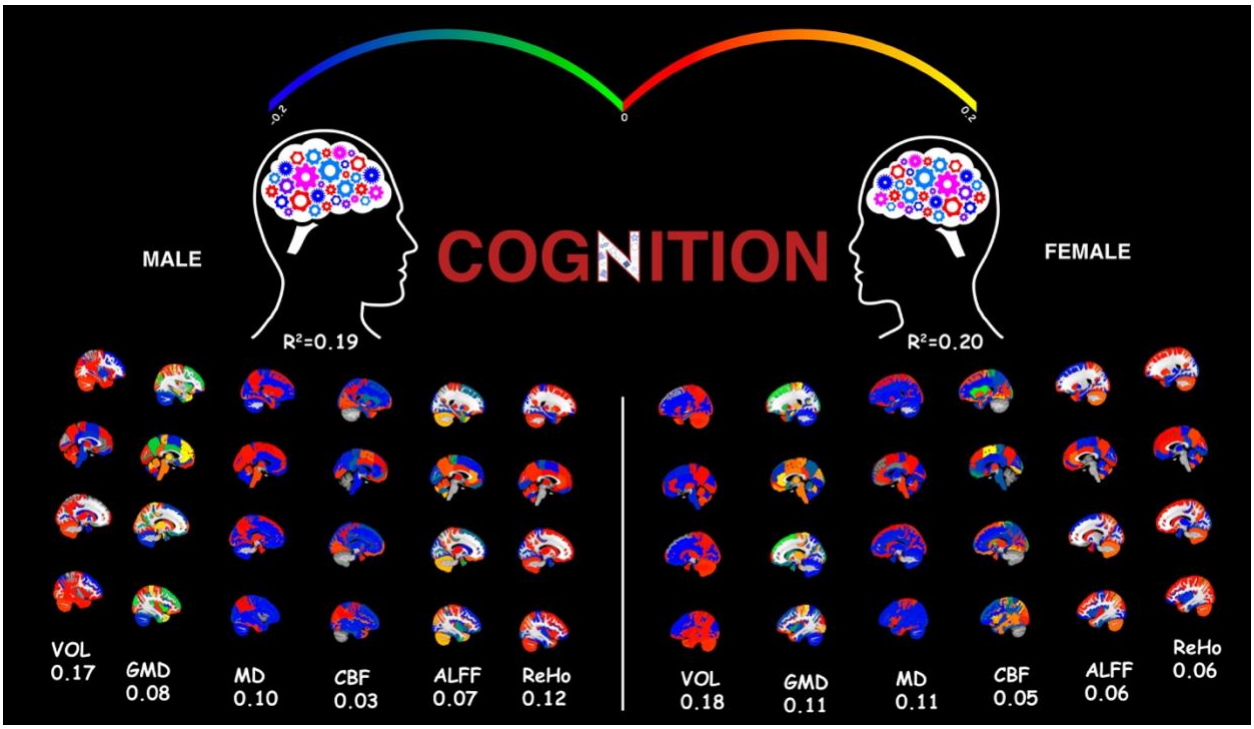


Figure 5. a. Results of ridge regressions predicting performance for each modality and across modalities. The distribution are shown of cross-validated R^2s for 1000 executions on real (blue bars) and permuted (red bars) data. b. Images showing regional correlations with performance for each parameter by sex.

bioRxiv preprint doi: https://doi.org/10.1101/659193; this version posted June 4, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

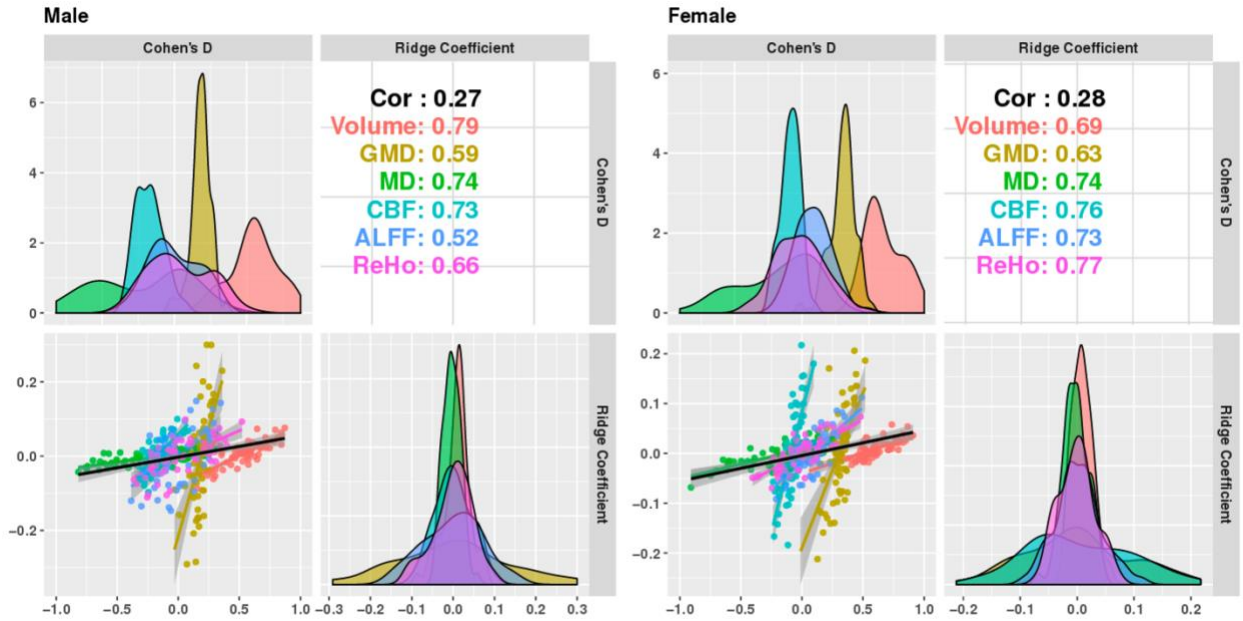


Figure 6. Scatterplot of effect sizes contrasting High and Low Performance groups with "Importance Parameters" from Ridge Regressions predicting performance for each modality and across modalities.