Supplemental Materials

A. Cohorts Used in Existing Studies

Only studies using resting-state functional connectivity for psychometric prediction in mainly healthy population with at least 200 subjects were included.

No.	Study reference			Cohort used
1	Sripada et al. 2019	Prediction of neurocognition in youth from resting state fMRI	Molecular psychiatry	ABCD
2	Greene et al. 2018	Task-induced brain state manipulation improves prediction of individual traits	Nature Communications	HCP-YA, PNC
3	Noble et al. 2017	Influences on test-retest reliability of functional connectivity MRI and its relationship with behavioral utility	Cerebral Cortex	HCP-YA
4	Yoo et al. 2019	Multivariate approaches improve the reliability and validity of functional connectivity and prediction of individual behaviors	NeuroImage	HCP-YA, recruited subjects
5	Finn et al. 2015	Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity	Nature Neuroscience	HCP-YA
6	Elliott et al. 2019	General functional connectivity: Shared features of resting-state and task fMRI drive reliable and heritable individual differences in functional brain networks	NeuroImage	HCP-YA, Dunedin Study
7	Sripada et al. 2019	Basic units of inter-individual variation in resting state connectomes	Scientific Reports	HCP-YA
8	Chen et al. 2019	Resting-state brain information flow predicts cognitive flexibility in humans	Scientific Reports	HCP-YA
9	Beaty et al. 2018	Robust prediction of individual creative ability from brain functional connectivity	PNAS	Recruited subjects
10	Jiang et al. 2020	Gender differences in connectome-based predictions of individualized intelligence quotient and sub-domain scores	Cerebral Cortex	UESTC, HCP-YA, COBRE
11	Ferguson et al. 2017	Fluid and flexible minds: Intelligence reflects synchrony in the brain's intrinsic network architecture	Network Neuroscience	HCP-YA
12	Jiang et al. 2018	Connectome-based individualized prediction of temperament trait scores	NeuroImage	UESTC, HCP-YA
13	Gao et al. 2019	Combining multiple connectomes improves predictive modelling of phenotypic measures	NeuroImage	HCP-YA, PNC
14	Kong et al. 2018	Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion	Cerebral Cortex	GSP, CoRR-HNU, HCP-YA
15	Li et al. 2019	Global signal regression strengthens association between resting-state functional connectivity and behavior	NeuroImage	GSP, HCP-YA
16	Dubois et al. 2018	A distributed brain network predicts general intelligence from resting-state human neuroimaging data	Philosophical Transactions B	HCP-YA
17	Kashyap et al. 2019	Individual-specific fMRI-subspaces improve functional connectivity prediction of behavior	NeuroImage	HCP-YA
18	Dubois et al. 2018	Resting-state functional connectivity best predicts the personality dimension of openness to experience	Personality Neuroscience	HCP-YA

19	Cui et al. 2018	The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features	NeuroImage	HCP-YA
20	Nostro et al. 2018	Predicting personality from network-based resting-state functional connectivity	Brain Structure and Function	HCP-YA
21	Li et al. 2019	Performing group-level functional image analyses based on homologous functional regions mapped in individuals	PLoS Biology	HCP-YA
22	Qin et al. 2019	Dissociating individual connectome traits using low-rank learning	Brain Research	HCP-YA
23	Jiang et al. 2020	Multimodal data revealed different neurobiological correlates of intelligence between male and females	Brain Imaging and Behavior	UESTC
24	Boeke et al. 2020	Toward robust anxiety biomarkers: A machine learning approach in a large-scale sample	Biological Psychiatry: Cognitive Neuroscience and Neuroimaging	GSP
25	Liu et al. 2018	Neural and genetic determinants of creativity	NeuroImage	Recruited subjects
26	He et al. 2020	Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics	NeuroImage	HCP-YA, UKB
27	Avery et al. 2020	Distributed patterns of functional connectivity predict working memory performance in novel healthy and memory-impaired individuals	Journal of Cognitive Neuroscience	HCP-YA
28	Burr et al. 2020	Functional connectivity predicts the dispositional use of expressive suppression but not cognitive reappraisal	Brain and Behavior	Recruited subjects
29	Maglanoc et al. 2019	Brain connectome mapping of complex human traits and their polygeninc architecture using machine learning	Biological Psychiatry	UKB
30	Cai et al. 2020	Brain functional connectome-based prediction of individual decision impulsivity	Cortex	HCP-YA
31	Takagi et al. 2019	State-unspecific patterns of whole-brain functional connectivity from resting and multiple task states predict stable individual traits	NeuroImage	HCP-YA
32	Ren et al. 2021	Connectome-based predictive modelling of creative anxiety	NeuroImage	Recruited subjects
33	Zhang et a. 2020	Do intrinsic brain functional networks predict working memory from childhood to adulthood?	Human Brain Mapping	GUSTO, CBDC, PING, BCAS
34	He et al. 2021	Executive function-related functional connectomes predict intellectual abilities	Intelligence	HCP-YA
35	He et al. 2021	Functional connectome prediction of anxiety related to the COVID-19 pandemic	The American Journal of Psychiatry	Recruited subjects
36	Tomasi and Volkow 2020	Network connectivity predicts language processing in healthy adults	Human Brain Mapping	HCP-YA
37	Sen et al. 2020	Predicting biological gender and intelligence from fMRI via dynamic functional connectivity	IEEE Transactions on Biomedical Engineering	HCP-YA
38	Liu et al. 2021	The functional connectome predicts feeling of stress on regular days and during the COVID-19 pandemic	Neurobiology of Stress	GBB, recruited subjects

ABCD = Adolescent Brain Cognitive Development; HCP-YA = Human Connectome Project Young Adult; PNC = Philadelphia Neurodevelopmental Cohort; UESTC = University of Electronic Science and Technology of China; COBRE = Center for Biomedical Research Excellence; GSP = Genomic Superstruct Project; CoRR-HNU = Consortium for Reliability and Reproducibility – Hangzhou Normal University; UKB = UK Biobank; GUSTO = Growing Up in Singapore Towards healthy Outcomes; CBDC = Cognition and Brain Development in Children; PING = Pediatric Imaging, Neurocognition, and Genetics; BCAS = Brain and Cognition Aging Study; GBB = Gene-Brain-Behavior

B. **fMRIPrep Methods**

The same pipeline was run for each subject separately, processing both its anatomical T1 image and all BOLD images. The boilerplate for the first subject is shown below.

```
Results included in this manuscript come from preprocessing
performed using *fMRIPrep* 20.1.1
(@fmriprep1; @fmriprep2; RRID:SCR_016216),
which is based on *Nipype* 1.5.0
(@nipype1; @nipype2; RRID:SCR_002502).

Anatomical data preprocessing

: A total of 2 T1-weighted (T1w) images were found within the input
BIDS dataset.
All of them were corrected for intensity non-uniformity (INU)
with `N4BiasFieldCorrection` [@n4], distributed with ANTs 2.2.0 [@ants,
RRID:SCR_004757].
The T1w-reference was then skull-stripped with a *Nipype* implementation of
the `antsBrainExtraction.sh` workflow (from ANTs), using OASIS30ANTs
as target template.
Brain tissue segmentation of cerebrospinal fluid (CSF),
white-matter (WM) and gray-matter (GM) was performed on
the brain-extracted T1w using `fast` [FSL 5.0.9, RRID:SCR_002823,
@fsl_fast].
A T1w-reference map was computed after registration of
2 T1w images (after INU-correction) using
`mri_robust_template` [FreeSurfer 6.0.1, @fs_template].
Brain surfaces were reconstructed using `recon-all` [FreeSurfer 6.0.1,
RRID:SCR_001847, @fs_reconall], and the brain mask estimated
previously was refined with a custom variation of the method to reconcile
ANTs-derived and FreeSurfer-derived segmentations of the cortical
gray-matter of Mindboggle [RRID:SCR_002438, @mindboggle].
Volume-based spatial normalization to two standard spaces (MNI152NLin6Asym,
MNI152NLin2009cAsym) was performed through
nonlinear registration with `antsRegistration` (ANTs 2.2.0),
using brain-extracted versions of both T1w reference and the T1w template.
The following templates were selected for spatial normalization:
```

FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model [@mni152nlin6asym, RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym], *ICBM 152 Nonlinear Asymmetrical template version 2009c* [@mni152nlin2009casym, RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym],

Functional data preprocessing

: For each of the 13 BOLD runs found per subject (across all
tasks and sessions), the following preprocessing was performed.
First, a reference volume and its skull-stripped version were generated
using a custom methodology of *fMRIPrep*.
Head-motion parameters with respect to the BOLD reference
(transformation matrices, and six corresponding rotation and translation
parameters) are estimated before any spatiotemporal filtering using
`mcflirt` [FSL 5.0.9, @mcflirt].
BOLD runs were slice-time corrected using `3dTshift` from
AFNI 20160207 [@afni, RRID:SCR_005927].
Susceptibility distortion correction (SDC) was omitted.
The BOLD reference was then co-registered to the T1w reference using
`bbregister` (FreeSurfer) which implements boundary-based registration [@bbr].
Co-registration was configured with six degrees of freedom.
The BOLD time-series were resampled onto the following surfaces
(FreeSurfer reconstruction nomenclature):
fsaverage.
The BOLD time-series (including slice-timing correction when applied)
were resampled onto their original, native space by applying
the transforms to correct for head-motion.
These resampled BOLD time-series will be referred to as *preprocessed
BOLD in original space*, or just *preprocessed BOLD*.
The BOLD time-series were resampled into standard space,
generating a *preprocessed BOLD run in MNI152NLin6Asym space*.
First, a reference volume and its skull-stripped version were generated
using a custom methodology of *fMRIPrep*.
Grayordinates files [@hcppipelines] containing 91k samples were also
generated using the highest-resolution ``fsaverage`` as intermediate
standardized
surface space.
Automatic removal of motion artifacts using independent component analysis
[ICA-AROMA, @aroma] was performed on the *preprocessed BOLD on MNI space*
time-series after removal of non-steady state volumes and spatial smoothing
with an isotropic, Gaussian kernel of 6mm FWHM (full-width half-maximum).
Corresponding "non-aggresively" denoised runs were produced after such
smoothing.
Additionally, the "aggressive" noise-regressors were collected and placed
in the corresponding confounds file.
Several confounding time-series were calculated based on the
preprocessed BOLD: framewise displacement (FD), DVARS and

three region-wise global signals.
FD was computed using two formulations following Power (absolute sum of relative motions, @power_fd_dvars) and Jenkinson (relative root mean square displacement between affines, @mcflirt).
FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* [following the definitions by @power_fd_dvars].
The three global signals are extracted within the CSF, the WM, and the whole-brain masks.
Additionally, a set of physiological regressors were extracted to allow for component-based noise correction [*CompCor*, @compcor].
Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor).
tCompCor components are then calculated from the top 5% variable voxels within a mask covering the subcortical regions.
This subcortical mask is obtained by heavily eroding the brain mask, which ensures it does not include cortical GM regions.
For aCompCor, components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). Components are also calculated separately within the WM and CSF masks.
For each CompCor decomposition, the *k* components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration.
The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file.
The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each [*@confounds_satterthwaite_2013*].
Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers.
All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces).
Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs),
configured with Lanczos interpolation to minimize the smoothing effects of other kernels [*@lanczos*].
Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.6.2 [@nilearn, RRID:SCR_001362],

```
mostly within the functional processing workflow.
For more details of the pipeline, see [the section corresponding
to workflows in *fMRIPrep*'s
documentation](https://fmriprep.readthedocs.io/en/latest/workflows.html
"FMRIPrep's documentation").


### Copyright Waiver

The above boilerplate text was automatically generated by fMRIPrep
with the express intention that users should copy and paste this
text into their manuscripts *unchanged*.
It is released under the
[CC0](https://creativecommons.org/publicdomain/zero/1.0/) license.

### References
```

C. **Supplementary Figure**

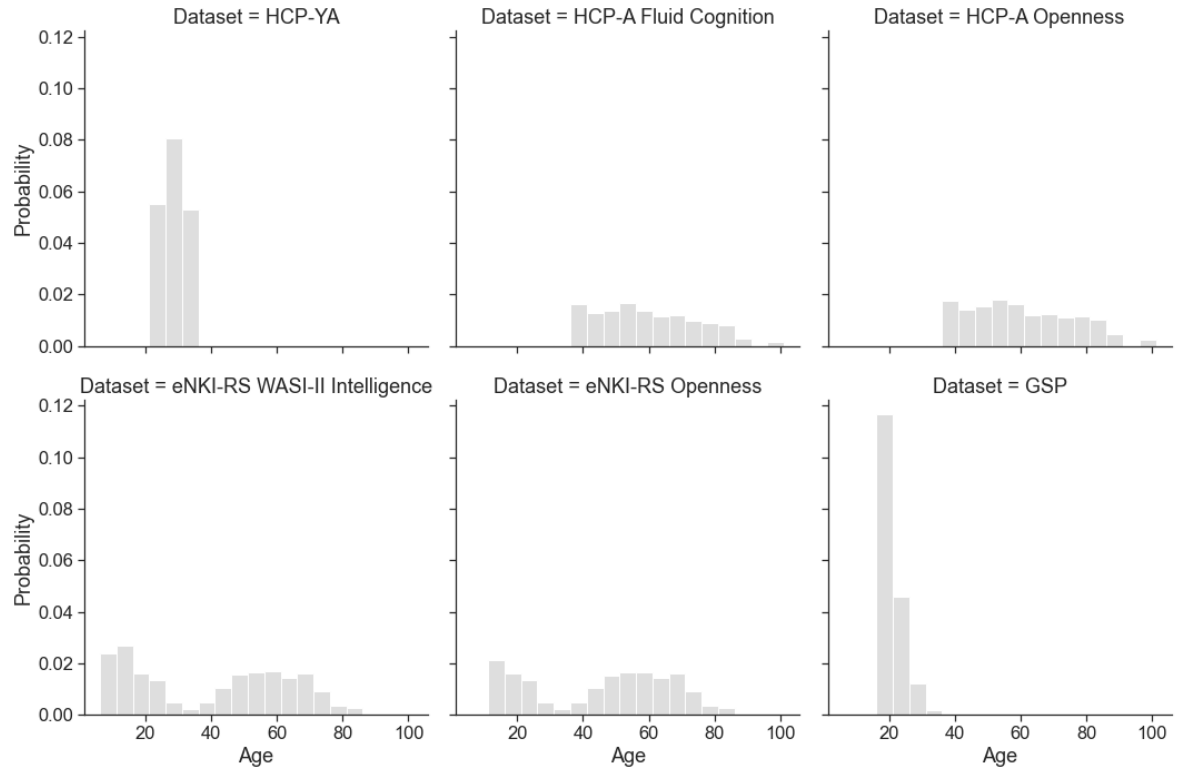

Figure S1. Distribution of age in the four cohorts. The distribution is plotted separately for fluid cognition and openness in HCP-A and eNKI-RS, as different measures are available in different subsets of the dataset in these two cases.

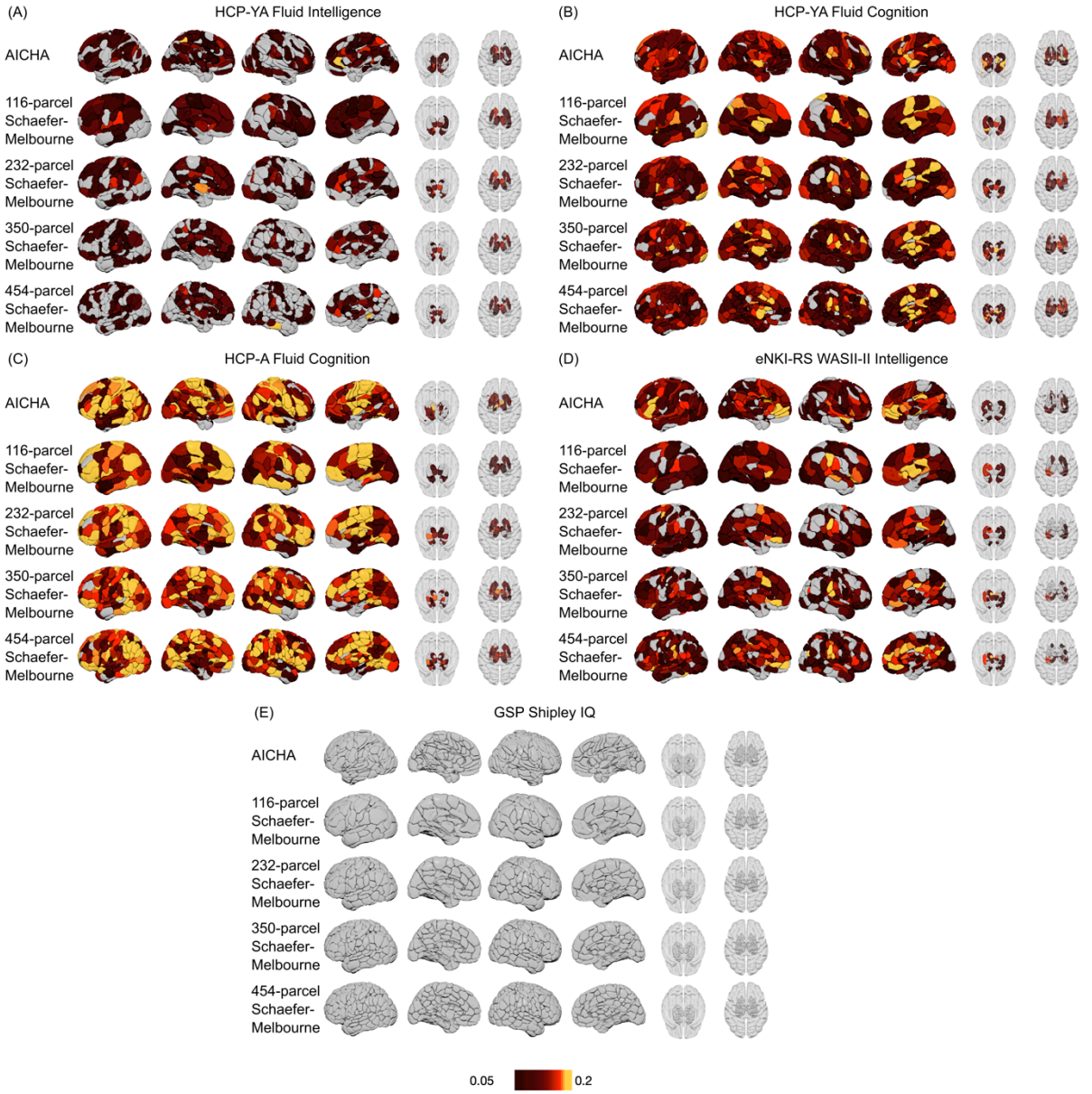

Figure S2. Prediction accuracy distribution maps of (A) HCP-YA fluid intelligence, (B) HCP-YA fluid cognition, (C) HCP-A fluid cognition, (D) eNKI-RS WASI-II intelligence, and (E) GSP Shipley IQ, using elastic net. Within each section, each row shows the prediction accuracy distribution overlaid on a parcellation used in region-wise CBPP prediction, in lateral and medial views of the left and right cortical hemispheres, as well as the bottom and top views of the subcortical regions. Color represents the magnitude of the prediction accuracies (Pearson correlation between predicted and observed values). Accuracies below 0.05 and non-significant accuracies are shown in gray.

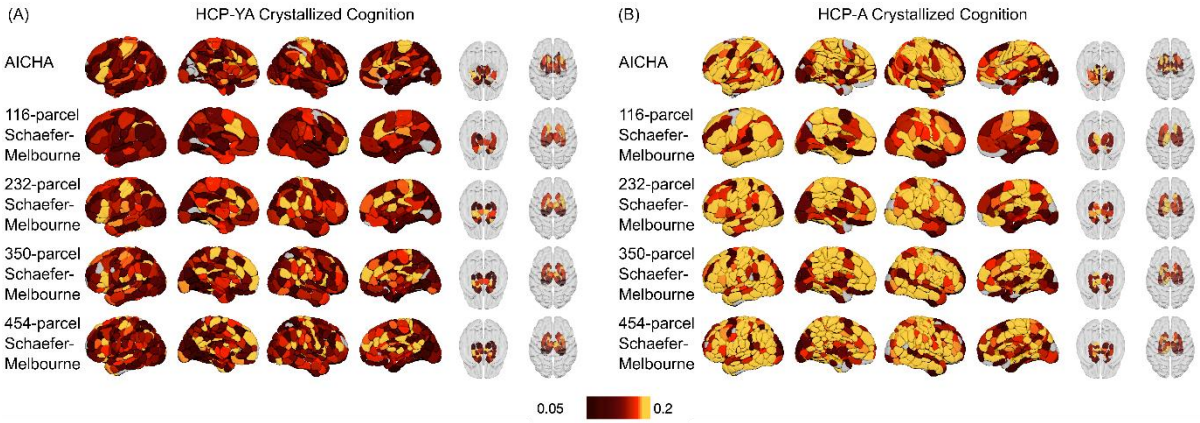

Figure S3. Prediction accuracy distribution maps of (A) HCP-YA crystallized cognition, and (B) HCP-A crystallized cognition. Each row shows the prediction accuracy distribution overlaid on a parcellation used in region-wise CBPP prediction, in lateral and medial views of the left and right cortical hemispheres, as well as the bottom and top views of the subcortical regions. Color represents the magnitude of the prediction accuracies (Pearson correlation between predicted and observed values). Accuracies below 0.05 and non-significant accuracies are shown in gray.

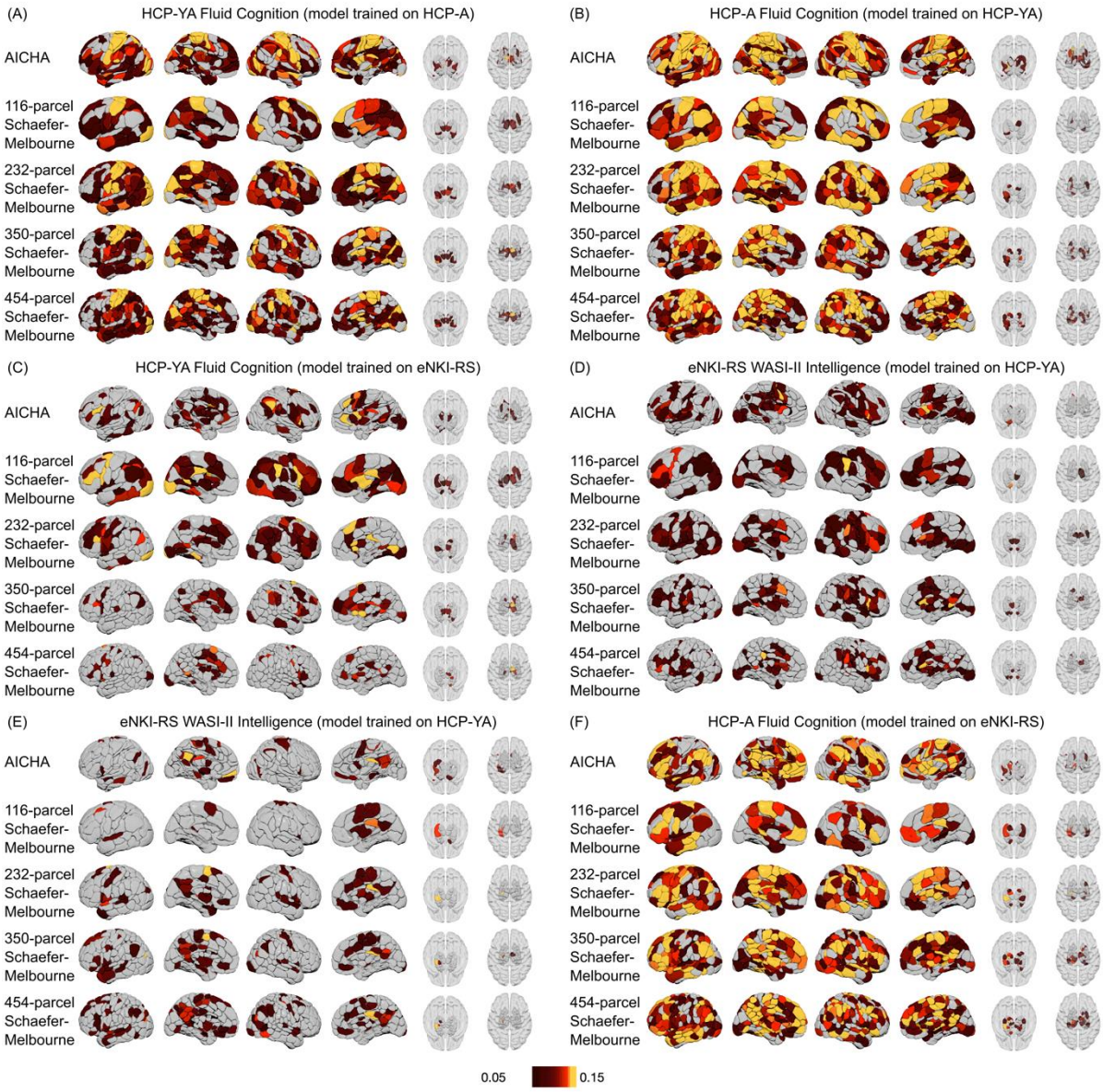

Figure S4. Prediction accuracy distribution maps of (A) HCP-YA fluid cognition with models trained on HCP-A data, (B) HCP-A fluid cognition with models trained on HCP-YA data, (C) HCP-YA fluid cognition with models trained on eNKI-RS data, (D) eNKI-RS WASI-II intelligence with models trained on HCP-YA data, (E) eNKI-RS WASI-II intelligence with models trained on HCP-A data, (F) HCP-A fluid cognition with models trained on eNKI-RS data. Each row shows the prediction accuracy distribution overlaid on a parcellation used in region-wise CBPP prediction, in lateral and medial views of the left and right cortical hemispheres, as well as the bottom and top views of the subcortical regions. Color represents the magnitude of the prediction accuracies (Pearson correlation between predicted and observed values). Accuracies below 0.05 are shown in gray.

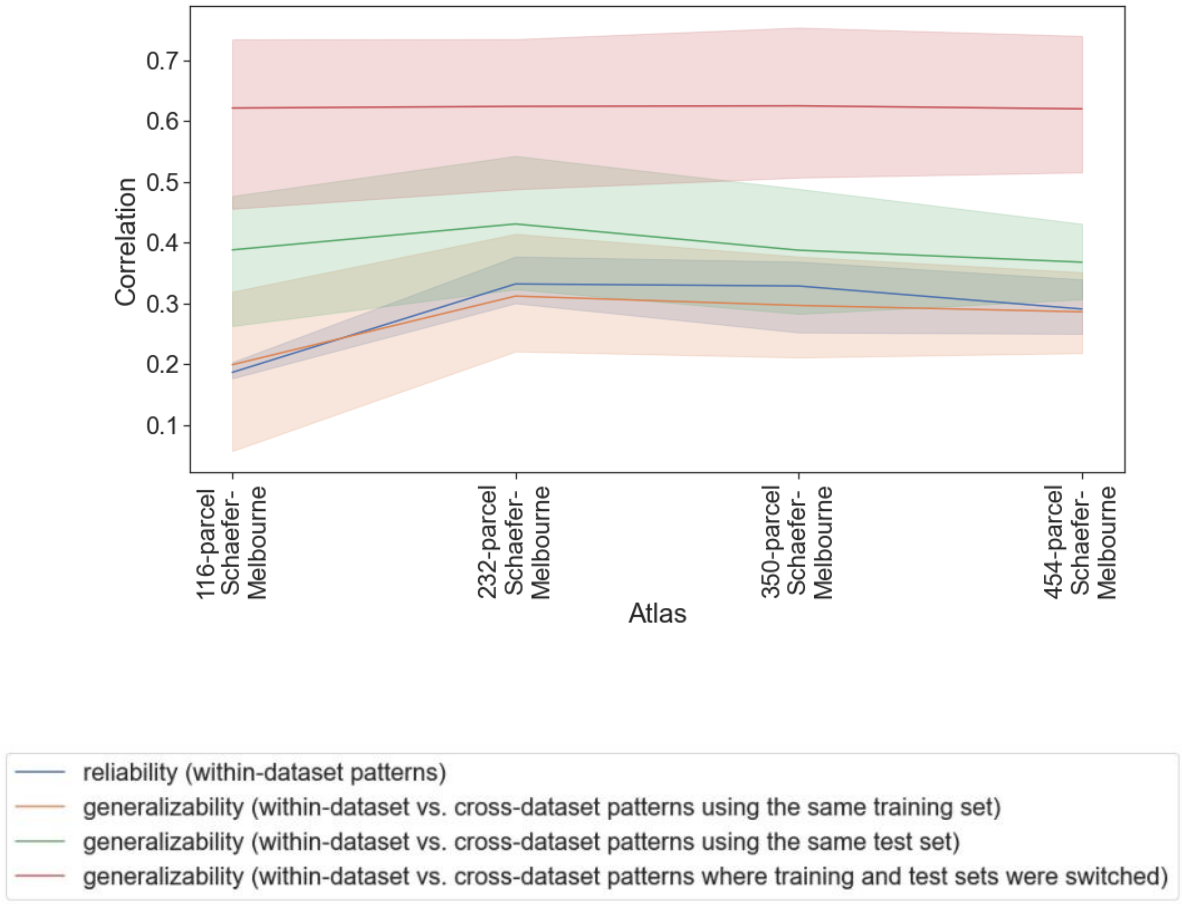

Figure S5. Correlation between prediction accuracy distributions of within-dataset prediction patterns (reliability) or between within-dataset and cross-dataset prediction patterns (generalizability), plotted against the Schaefer-Melbourne atlas with increasing granularity. Each line shows the mean correlation values across all pairs of comparisons, while the colored areas around the lines represent the 95% confidence interval.