PNAS
www.pnas.org

Supplementary Information for

Neurobehavioural Correlates of Obesity are Largely Heritable

Uku Vainik, Travis Baker, Mahsa Dadar, Yashar Zeighami, Andréanne Michaud, Yu Zhang, José C. García Alanis, Bratislav Misic, D. Louis Collins, Alain Dagher

https://doi.org/10.1101/204917

Alain Dagher

Email: alain.dagher@mcgill.ca

This PDF file includes:

Supplementary text

Figs. S1 to S12

Table S1

References for SI reference citations

Other supplementary materials for this manuscript include the following:

Dataset S1

Supplementary Information Text

Methods

Participants

Data were provided by the Human Connectome Project (24) WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657, RRID:SCR_008749) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University,

The analyzed data were split between the S900 data release (964 participants) and the S1200 data release (236 additional participants). We treated the S900 as the main analysis sample and results from this sample are reported throughout the paper. At times, we used unique participants from the S1200 release for replication, referred to as S1200n. For the main analysis sample, we applied the following exclusion criteria, as these might confound brain-obesity associations: people with missing values on crucial variables, such as age, BMI, education, income, gender, race, and ethnicity (n=6), hypo/hyper thyroidism (n=4), other endocrine problems (n=16), underweight (BMI <=18, n=9), and women who had recently given birth (n=9). In addition, as we used family information to control for participants' relatedness, we excluded participants that were half-siblings to other participants (n=31). The same exclusions were applied to S1200n (n=11).

The final main analysis dataset consisted of 895 participants, demographics of which are summarized in Table 1. The sample had good gender balance and variation in BMI and income. As limitations, the sample was relatively young and well educated, and BMI distribution was slightly less obese compared to current prevalence estimates for Missouri or the US as a whole (MO: 31.7%, US: 36.5%, ref: , 49). Most people were white and non-Hispanic, however other races-ethnicities were also represented. The participants were nested into 384 families, typically having 1 to 3 siblings in the dataset. For comparison, we also provide the same statistics for the S1200n sample, as well as a subset of S1200n sample in which no participant is related to the S900 sample.

For the heritability analysis between each neurocognitive factor and BMI, we randomly chose one sibling pair per family, ensuring that the pair had complete data. Non-twin sibling pairs were considered equivalent to dizygotic twin pairs with respect to heritability analyses once data was

residualized for age and gender. If multiple sibling pairs within a family had complete data, we prioritized choosing monozygotic twin pairs and dizygotic twin pairs over non-twin sibling pairs. Depending on the neurocognitive factor, the heritability analysis was conducted on 46-111 pairs of monozygotic twins (median=97) and 60-202 pairs of dizygotic twins and siblings (median=176).

Measures

Psychological measures.

Participants completed an extensive set of questionnaires and cognitive tests (see 53, 54 for an overview). In the current analysis, we included 22 questionnaires and 18 cognitive tests (see Fig. 2 and Dataset S1, section 1 for complete list). Here we refer to the set of questionnaire results as personality variables, as personality encompasses various patterns of what people want, say, do, feel, or believe (55). Based on our previous review (6) we chose cognitive tests capturing aspects of executive function, memory, and language.

Cortical thickness.

All T1-weighted MRI images were processed using the CIVET pipeline (version 2.0) (29, 56, 57). Processing was executed on the Canadian Brain Imaging Network (CBRAIN) High Performance Computing platform for collaborative sharing and distributed processing of large MRI datasets (58). Briefly, native T1-weighted MRI scans were corrected for non-uniformity using the N3 algorithm (59). The corrected volumes were masked and registered into stereotaxic space, and then segmented into gray matter (GM), white matter (WM), cerebrospinal fluid (CSF) and background using a neural net classifier (60). The white matter and gray matter surfaces were extracted using the Constrained Laplacian-based Automated Segmentation with Proximities algorithm (61, 62). The resulting surfaces were resampled to a stereotaxic surface template to provide vertex based measures of cortical thickness (63). All resulting images were visually inspected for motion artefacts by experienced personnel and then subsequently processed through a stringent quality control protocol, which only 641 of the 894 participants in our initial cohort passed. In the S1200n, 144 of the 214 passed. For those participants who passed, cortical thickness was then measured in native space using the linked distance between the two surfaces across 81924 vertices and a 20mm surface smoothing kernel was applied to the data (64). The Desikan–Killiany–Tourville (DKT) atlas was used to parcellate the surface into 64 cortical

regions (65). Cortical thickness was averaged over all vertices in each region of interest for each subject (66) and the effect of mean cortical thickness was regressed to allow for regional analysis (67). After participant exclusions, data was available for 591/137 participants in the S900/S1200n samples.

Volumetric estimates.

Because the CIVET cortical thickness method does not cover all medial temporal and subcortical structures, we used volumetric estimates for these brain regions. For subcortical volumetric estimation, T1-weighted scans of the subjects were pre-processed through a computerized pipeline (n=899). Image denoising (68), intensity non-uniformity correction (59), and image intensity normalization into range (0-100) using histogram matching were performed. After preprocessing, all images were first linearly (using a 9-parameter rigid registration) and then nonlinearly registered to an average template (MNI ICBM152) as part of the ANIMAL software (30, 69). The subcortical structures, i.e., thalamus, putamen, caudate, and globus pallidus were segmented using ANIMAL by warping segmentations from ICBM152 back to each subject using the obtained nonlinear transformations. The medial temporal lobe structures, i.e. hippocampus, amygdala, temporal pole, and parahippocampal, entorhinal and perirhinal cortices, were segmented using an automated patch-based label-fusion technique (70). The method selects the most similar templates from a library of labelled MRI template images, and combines them with a majority voting scheme to assign the highest weighted label to every voxel to generate a discrete segmentation. Quality control was performed on the individual registered images as well as the automated structure segmentations by visual inspection, and inaccurate results were discarded. In S900, 648 participants passed the quality control for medial temporal lobe structures, ad 895 for subcortical structures. Within S1200n, of the 214 participants, 212 passed the quality control for subcortical structures, and 174 passed the quality control for medial temporal lobe. After exclusions, the S900/S1200n samples included data from n=828/204, 8 parcels per subjects for the subcortical structures, and n=594/166, 12 parcels for the medial temporal lobe structures.

Data Analysis

Analyzing each feature.

A schematic pipeline of the analysis is displayed in SI Appendix Fig. S1. Data from all

neurocognitive factors were first residualized for control variables (age, ethnicity, gender, handedness, race) using multiple linear regression. When presenting phenotypic associations, we used a linear mixed model, adding a random intercept for family (SI Appendix Fig. S1), and also varied the involvement of income and education. As BMI was skewed (long-tail at the upper end of the scale), it was log-transformed to achieve a normal-like distribution. Handedness was also log normalized.

For each factor category (cognition, personality, cortical thickness, medial temporal volume, subcortical volume), factor-BMI relationships were assessed using univariate correlation between each brain parcel or test score and BMI. We initially also tried using a partial least squares (PLS) correlation approach, which is a multivariate technique suited to handling correlated predictors (71, 72). However, the PLS estimates were extremely close to univariate correlations, therefore univariate correlations were preferred for simplicity. As a result, we received an estimate of the relative contribution (weight) of each predictor within a given factor. Estimates used in this study are presented in Dataset S1, section 2.

Creating poly-phenotype scores.

To summarize effects for each neurocognitive factor, we created an aggregate BMI risk score or *poly-phenotype score (PPS)* for each neurocognitive factor. This was inspired by the polygenic risk score approach, where the effects of single-nucleotide polymorphisms are added up to form a total genetic score (73). Specifically, we used the correlation-derived weights to multiply each participant's measured values, and aggregated the results into a single composite variable for a given factor, the PPS. A PPS would reflect the total association that a given factor has with BMI. Even though only some features within a neurobehavioural factor had significant effects on BMI, and certain features correlated with each other (see Datasets S3-S7), both our testing (see SI Results) and recommendations by others (74) lead us to not apply p-value cutoffs, clumping, or pruning, as excluding these steps does not hurt predictive ability and improves transparency (74). PPS-s have a mean of 0 but varying standard deviation, depending on the number of features and their effect sizes (Dataset S1, section 8).

We used cross-validation principles to avoid and test for overfitting. Namely, we divided participants into 10% folds. Each 10% fold received the correlation weights from the remaining

90% of the sample. As the result, we received one PPS vector for each factor, where each participant's score was based on out-of-sample prediction. When creating the 10% folds, we created folds for each factor separately, as each factor has a different number of available data points, ensuring that folds were as equal in size as possible. We also ensured that siblings from the same family were in the same fold. Therefore, no data from family members were used in calculating both the correlation weights and performing out of sample predictions.

To test the robustness of PPS-s, we first tested the impact of not pruning and applying p-value cutoffs. In a pruned PPS, features are omitted that a) correlate above criterion to another feature and b) have lower correlation with BMI than the other feature (75). In a PPS with p-value cut-off, features are omitted that have an above-criterion uncorrected p-value when correlated with BMI Neither pruning nor a p-value cutoff improved the predictive ability of the PPS-s (see SI Results).

We further tested the predictive ability of PPS scores by applying the weights created on the full S900 release to predict BMI in the S1200n release (new participants only), which we did not touch before predicting. As 101 participants within the S1200n were related to participants in the S900, we also tested the predictive ability in the subset of S1200 that was not related to S900 (n=124).

Heritability analysis.

In the heritability analysis, a typical behavioural genetics decomposition uses relatedness assumptions between individuals to divide variance in a trait to the following components: genetic variance (A, additive and interactive effects), shared environmental variance (C, family and shared school effects), and unique environmental variance (E, unique experience and measurement error). The assumptions are: 100% of genetic variance shared between monozygotic twins, 50% of genetic variance shared between dizygotic twins and sex-and gender residualized siblings, 100% of family environment shared by all siblings, 0% unique variance shared between siblings. Such decomposition is called univariate heritability.

Besides establishing univariate heritability, one can also conduct heritability analysis on the covariance between two traits. For instance, a genetic correlation is the correlation between the A components of trait 1 and trait 2. A bivariate heritability analysis decomposes the phenotypic

correlation between trait 1 and trait 2 into A, C, and E components.

Heritability analysis was conducted on PPS scores not residualized for family structure, as this information is used in heritability modelling. We then ran bivariate heritability analyses separately between each PPS and BMI, which provided univariate heritability estimates of the PPS-s and BMI, genetic and environmental correlations between the univariate estimates of PPS-s and BMI, and bivariate decomposition of the phenotypic correlation between each PPS and BMI. We used the AE model, since BMI was best explained by an AE model, as opposed to an ACE model, based on Akaike Information Criterion (AIC) (Dataset S1, section 9). Similar AIC patterns were present for bivariate models (SI Appendix Fig. S12, Dataset S1, section 12). We report only standardized A estimates in the main results, as in the univariate and bivariate analysis of the AE model, E=100-A. Also, no environmental correlations were significant. All standardized and unstandardized estimates are reported in the supplementary materials (Datasets S10-S11).

Analysis software.

Analysis was conducted in Microsoft R Open 3.4.0 (76), using May 2017 version of packages abind, car, caret, cowplot, corrplot, ggplot2, lme4, MuMIn, pbkrtest, plyr, psych, synthpop, tidyr, WriteXLS (77–92). Cortical thickness was plotted using Surfstat (93) in MATLAB (94). Heritability analysis was conducted using OpenMX (95), adapting scripts provided by the Colorado International Twin Workshop (96).

SI Results

Control variables

Age, gender and race related to BMI, demonstrating the need for residualizing (SI Appendix Fig. S2). Marginal R^2 explaining only fixed effects was 0.07, and conditional R^2 explaining both fixed and random effects was 0.38, highlighting the effect of family structure. When controlling for education and income, education was a significant additional predictor, with total model R^2 being 0.09 and conditional R^2 0.37. Further, controlling for family structure in a nested model as random intercept improved model fit (AIC dropped from 7006 to 6895 / 6978 to 6885 when controlling for education and income), suggesting that family nesting needs to be taken into account.

Robustness of PPS-s

Similarly to genetic literature (74), we found that pruning features or applying a p-value threshold does not change the predictive ability of the PPS-s (SI Appendix Fig. S5 & S6).

To test the generalizability of the PPS approach, we used weights obtained from the full S900 release (SI Appendix Fig. S3 right and S4 right) to predict the BMI of new participants in the S1200 release (S1200n, n=236), which were not used in any of the initial assessments. As certain participants in the S1200n release were related to participants in the S900, we also tested the PPS performance when they were excluded. As can be seen in SI Appendix Fig. S7, cortical thickness estimates are very similar, no matter the training or testing dataset. Cognition PPS effect sizes were similar to each other, but did not reach statistical significance in the replication sample (S1200n). Personality PPS had unexpectedly high correlation with BMI in the new data. Further research is needed to determine if such effect sizes would further replicate. Medial temporal lobe PPS-s also did not replicate.

Heritability replication

We tested whether the PPS-based bivariate analysis patterns would replicate in the S900 dataset, but using unaggregated top individual features within the PPS-s. We chose the 5 individual features from the top predictors of cognition and cortical thickness. As shown in SI Appendix Fig. S8, the individual tasks are comparable with the PPS-s in terms of univariate heritability, genetic correlations, and heritability of phenotypic correlation. However, with genetic correlations, the estimates are non-significant (SI Appendix Fig. S8 B1&B2), suggesting that we are not powered to establish significance of the smaller correlations. Further, the standardized estimates for heritability of the phenotypic correlations (SI Appendix Fig. S8 C1&C2) are noisier and the estimator often failed at estimating standardized confidence intervals. Such failures at individual feature levels highlight the value of PPS-s, which provide more stable estimates at these sample sizes.

We further used participants only in the S1200n release to replicate the bivariate heritability analysis results in new data. PPS weights were obtained from the S900 release. We focused only on participants who did not have siblings in the S900 release. Granted, the power is low because of fewer complete twin pairs available (29 MZ pairs and 30 DZ pairs). The univariate estimate for BMI heritability was [A=64% [95% CI: 41%;79%]. In the bivariate analysis, we were also

able to replicate the patterns seen in the main dataset (SI Appendix Fig. S9), however the confidence intervals were often covering 0 or not estimated, likely due to small sample size.

Figures

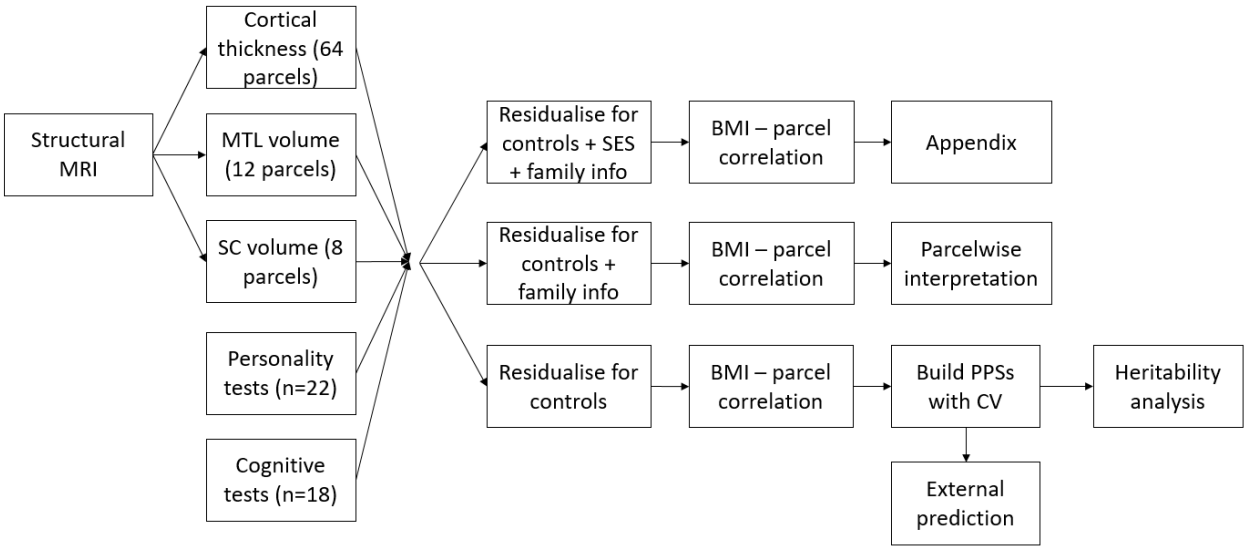

Fig. S1. A schematic diagram of the analysis pipeline. All steps were conducted on all neurocognitive factors separately. BMI=body mass index; CV=cross-validation; MTL=medial temporal lobe; MRI=magnet resonance image; PPS=poly-phenotype score; SC=subcortical; SES=socio-economic status (education and income).

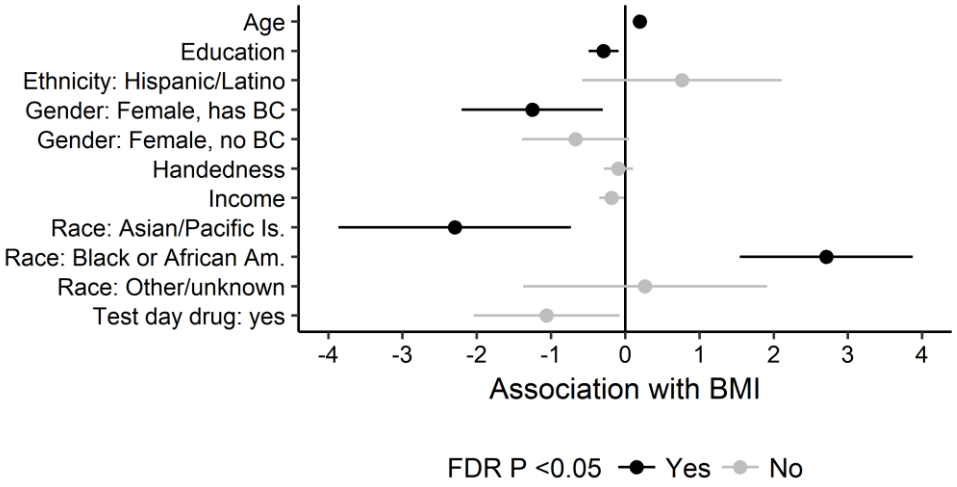

Fig. S2. Regression weights of a multilevel linear model nested for family. Lines mark standard 95% confidence intervals. Intercept is 27.37 (standard error: 2.16). For interpretability, regular BMI is unscaled here. Reference groups: Gender: male, Race: white, Ethnicity: not Hispanic/unknown. Am.=American; BC=birth control; Is.=Islander

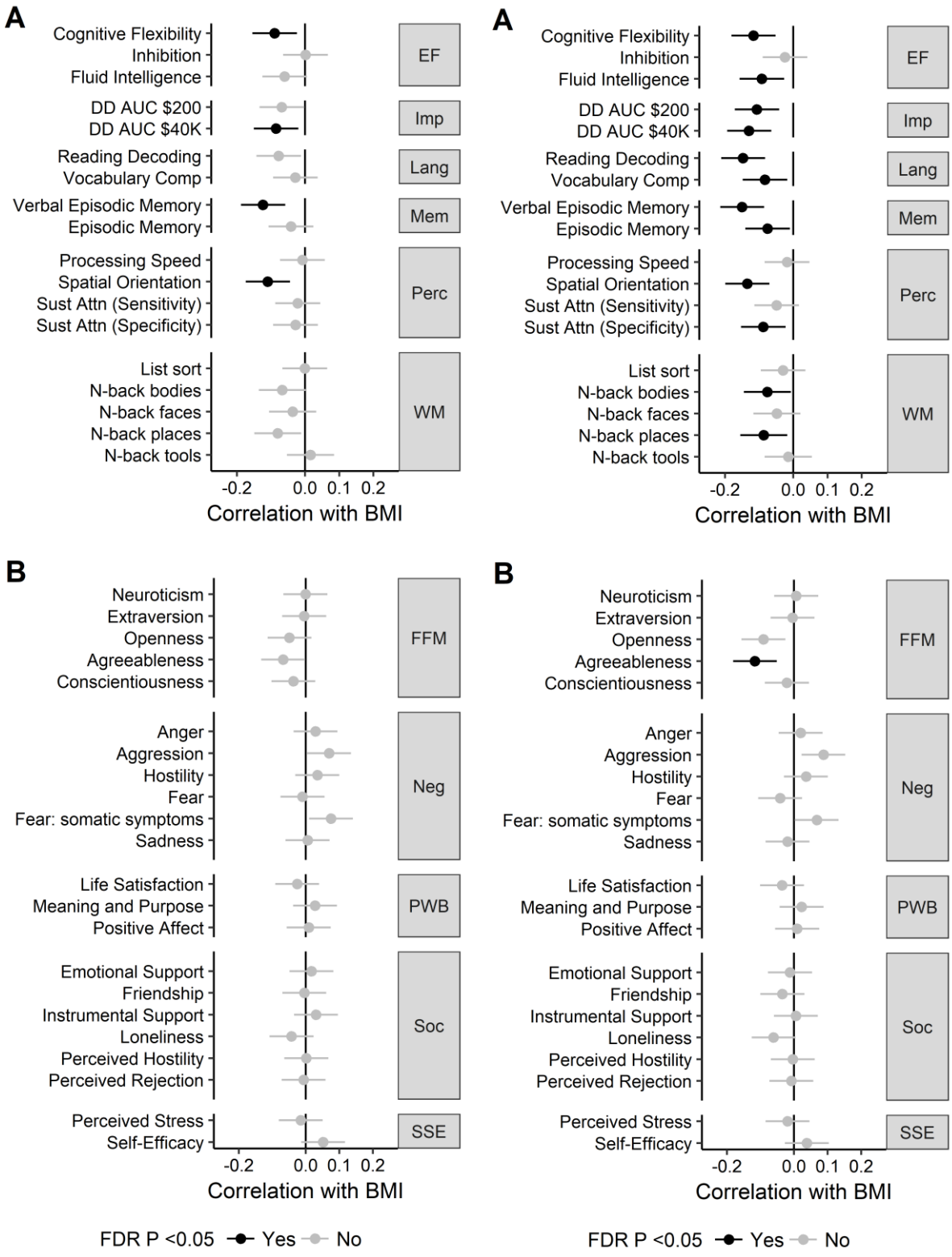

Fig. S3. Associations between body mass index (BMI), cognitive test scores (A), and personality

traits (B), either when controlling for education, income, and family structure (left), or not controlling for these variables (right). Error bars mark 95% confidence intervals. See Dataset S1, section 1 for explanation of cognitive test names. Numerical values are reported in Dataset S1, section 2. EF=executive function; FFM=Five-Factor Model; FDR=false discovery rate; Imp=(lack of) impulsivity; Lang=language; Mem=memory; Neg=negative affect; Perc=perception; PWB=psychological well-being; Soc=social relationships; SSE=stress and self efficacy; WM=working memory

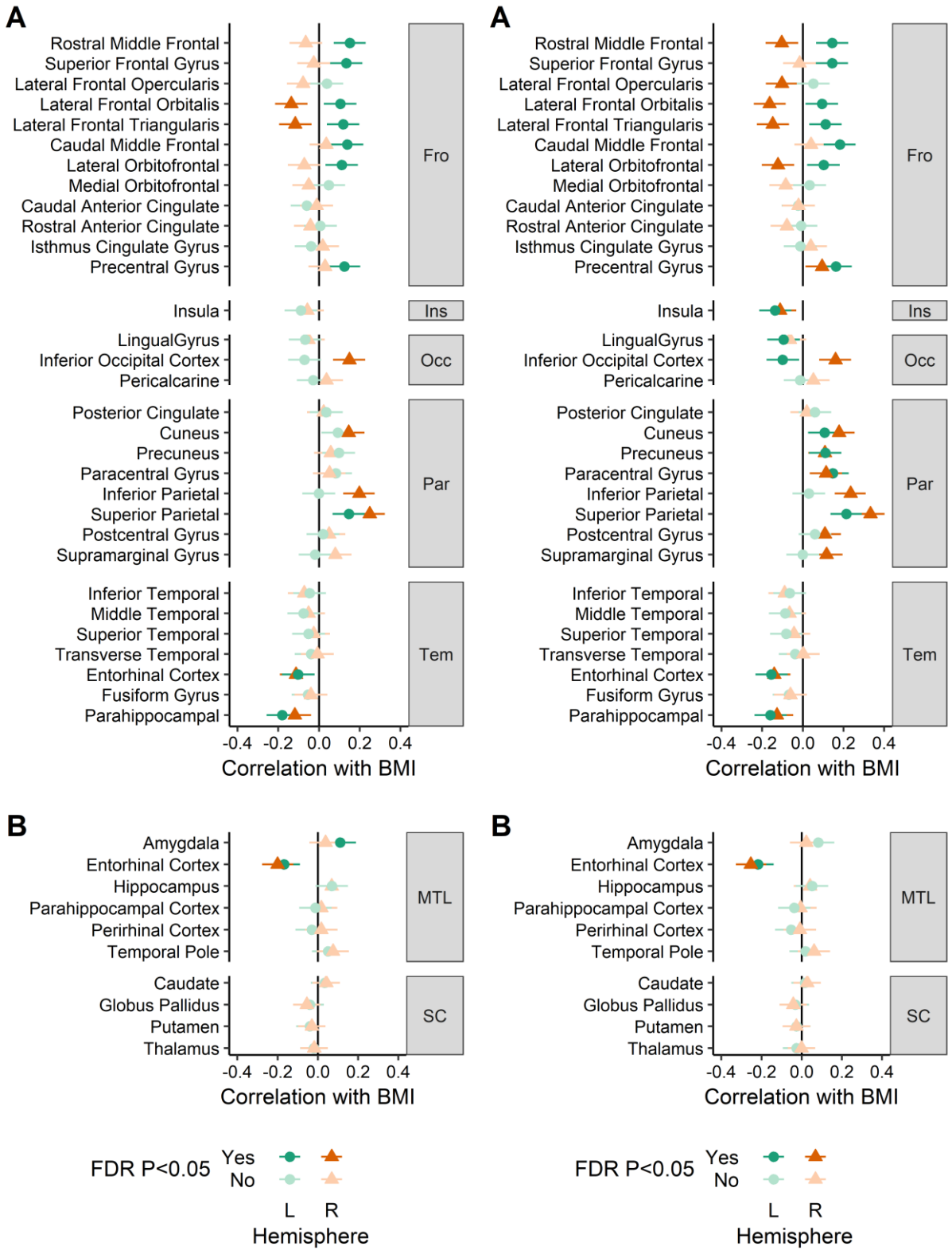

Fig. S4. Associations between body mass index (BMI), cortical thickness (A) and regional brain

volume (B), either when controlling for education, income, and family structure (left), or not controlling for these variables (right). Error bars mark 95% confidence intervals. Numerical values are reported in Dataset S1, section 2. FDR=false discovery rate; Fro=frontal, Ins=insula; L=left; Occ=occipital; Par=parietal; R=right; Tem=temporal;

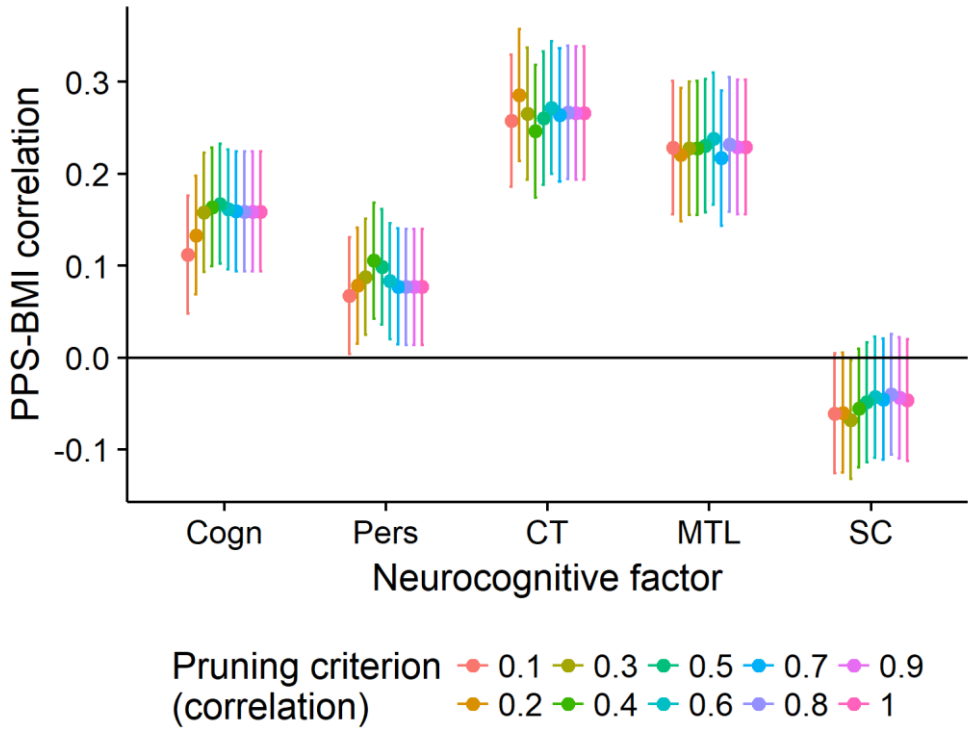

Fig. S5. Low impact of pruning to the poly-phenotype scores' (PPS) associations with BMI. PPS-s were trained and tested within the Human Connectome Project's S900 release, using cross-validation. Pruning means excluding features that have a higher correlation than set criterion with another feature that associates with BMI. A pruning criterion equal to 1 means no pruning was done. Cogn=PPS of cognitive tests; CT=PPS of cortical thickness; MTL=PPS of medial temporal

lobe volume; Pers=PPS of personality tests.

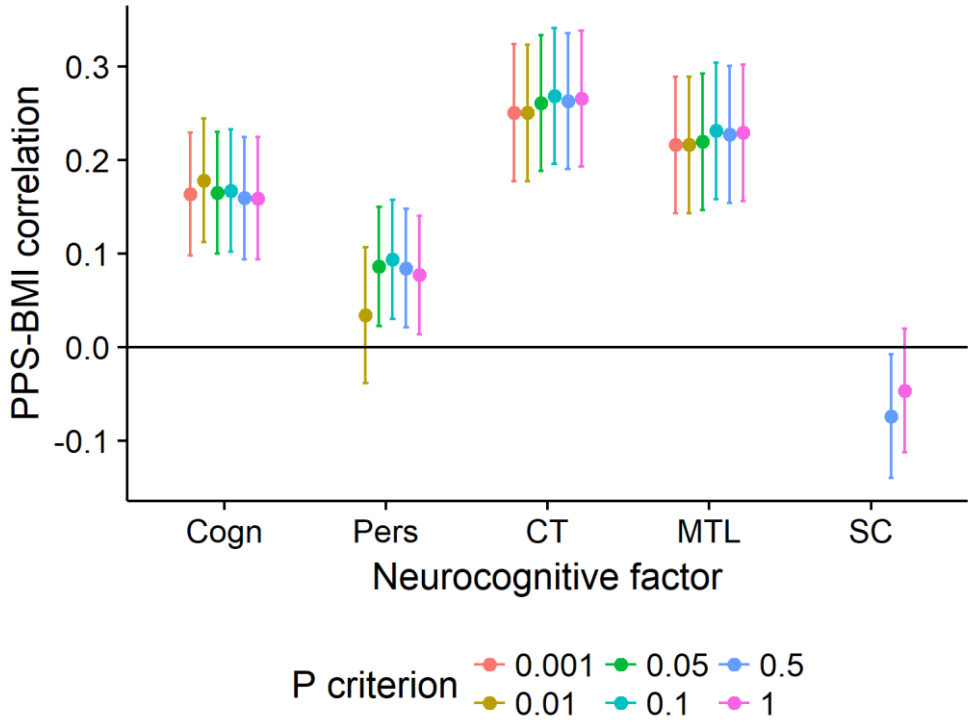

Fig. S6. Low impact of excluding features by p value to the poly-phenotype scores' (PPS) associations with BMI. PPS-s were trained and tested within the Human Connectome Project's S900 release, using cross-validation. Features with a p value higher than criterion were excluded from the PPS. A p criterion of 1 means no exclusion was done. Cogn=PPS of cognitive tests; CT=PPS of cortical thickness; MTL=PPS of medial temporal lobe volume; Pers=PPS of personality tests.

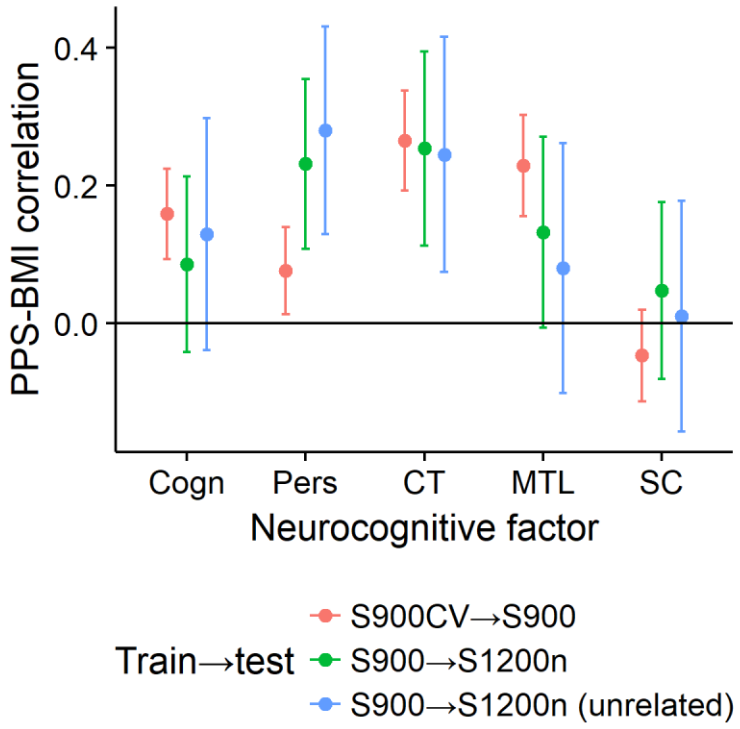

Fig. S7. Comparison of poly-phenotype scores' (PPS) performance in correlating with BMI, depending on training data and test data.

S900CV→S900: PPS-s within S900 release trained and tested with cross-validation to avoid bias. These PPS-s are used in heritability analysis.

S900→S1200n: PPS-s trained on S900 and tested in full S1200n sample.

S900→S1200n (unrelated): PPS-s trained on S900 and tested in S1200n sample not related to S900.

Cogn=PPS of cognitive tests; CT=PPS of cortical thickness; CV=cross-validated; MTL=PPS of medial temporal lobe volume; Pers=PPS of personality tests; S900 – Participants in Human Connectome Project's S900 release; S1200n – participants only in the S1200 release; SC=PPS of subcortical structure volumes;

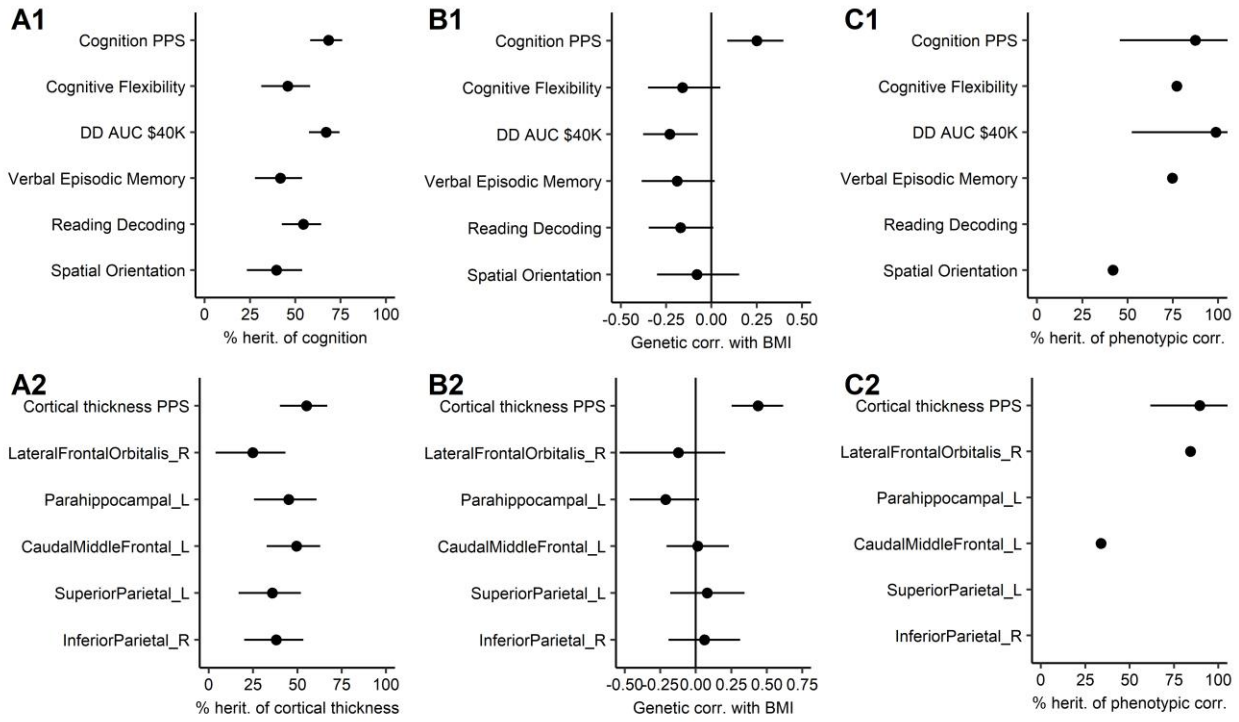

Fig. S8. Heritability analysis of the association between poly-phenotype scores (PPS) of cognitive test scores (A1-C1) and cortical thickness (A2-C2), compared with most significant individual features of each PPS. (A) Heritability of each trait. The effect of unique environment (E) is not shown, since E=100-A. (B) Genetic correlations between BMI and each PPS or between BMI and each feature. The PPS-based genetic correlations are positive, because the PPS-s are designed to positively predict BMI. However, individual features can have negative genetic correlations. (C) Heritability of the phenotypic correlation between BMI and PPS or between BMI and each feature. Horizontal lines depict 95% confidence intervals. The estimator failed at estimating certain features. Corr=correlation; L=Left hemisphere; herit=heritability; R=right hemisphere.

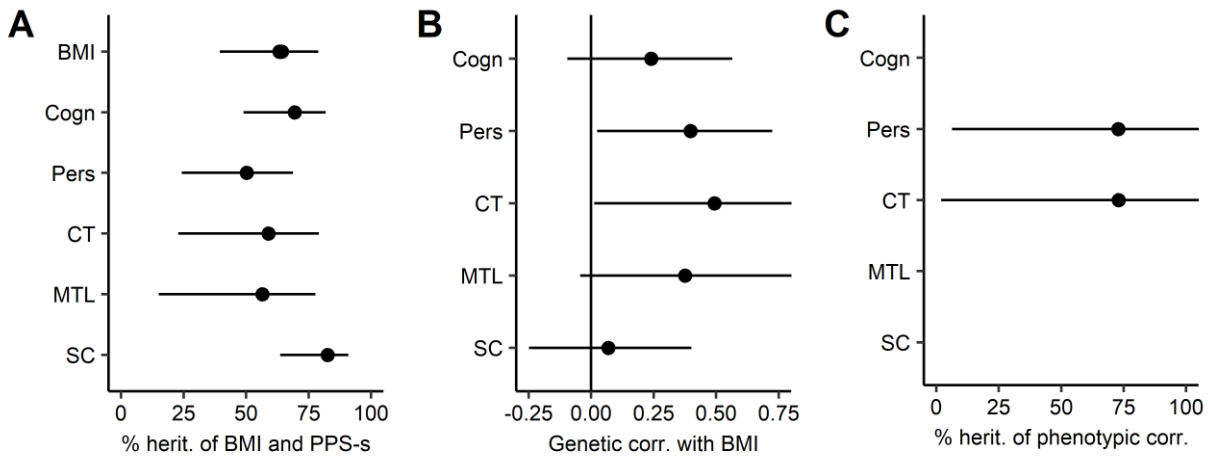

Fig. S9. Heritability analysis of the association between poly-phenotype scores (PPS) and body mass index (BMI) in the S1200n sample unrelated to S900. (A) Heritability of each trait. BMI has multiple estimates, since it was entered into a bivariate analysis with each PPS separately. The effect of unique environment (E) is not shown, since E=100-A. (B) Genetic correlations between BMI and each PPS. The genetic correlations are positive, because the PPS-s are designed to positively predict BMI. None of the environmental correlations were significant and therefore not shown. (C) Heritability of the phenotypic correlation between BMI and PPS. Horizontal lines depict 95% confidence intervals. Estimates not shown for PPS-s that did not have significant phenotypic association with BMI. Cogn=PPS of cognitive tests; corr=correlation; CT=PPS of cortical thickness; herit=heritability; MTL=PPS of medial temporal lobe volume; Pers=PPS of personality tests; SC=PPS of subcortical structure volumes

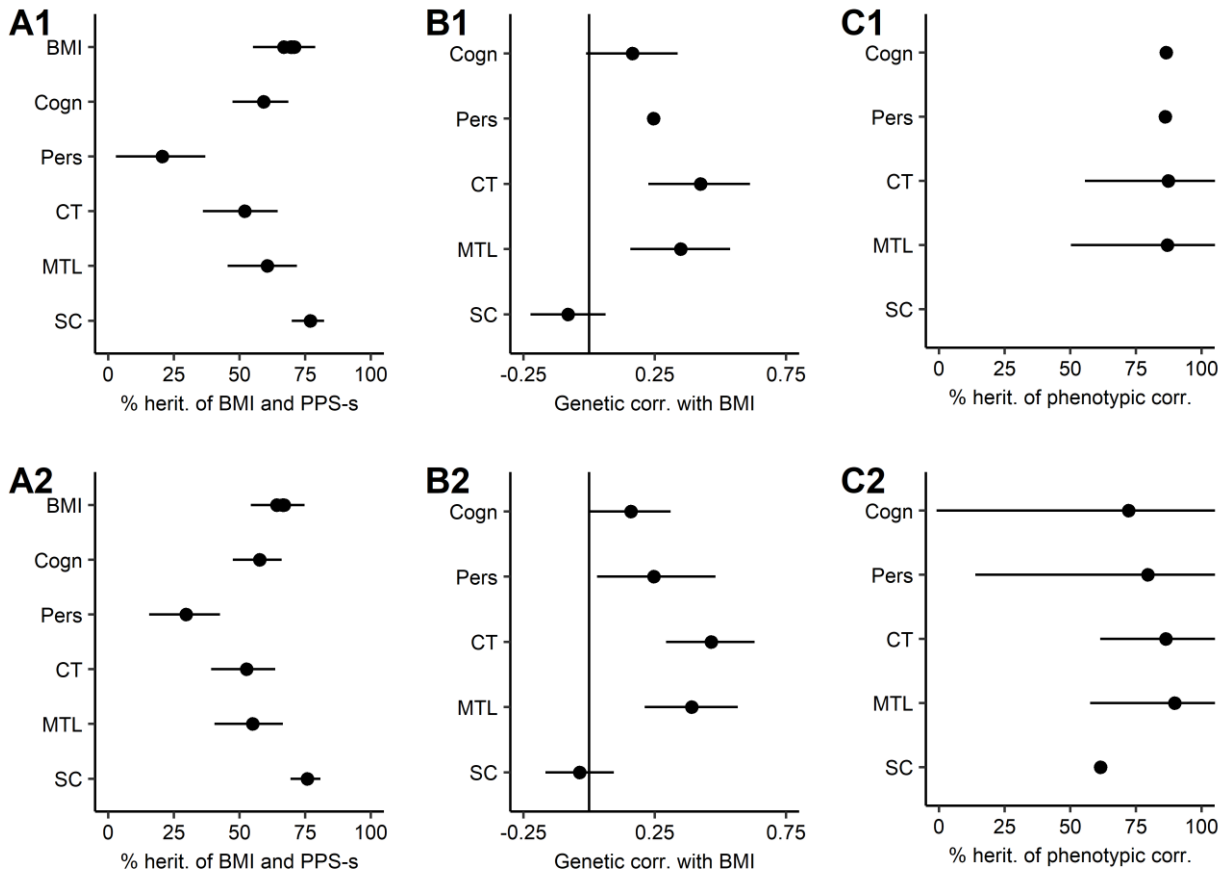

Fig S10. Heritability analysis of the association between poly-phenotype scores (PPS) and body mass index (BMI), when controlling for education and income within the S900 sample (top panel) and in the S1200 sample, where S1200n is added to the S900 sample (bottom panel). As in previous analyses, the PPS weights of S1200n sample are based on S900 sample, S1200n sample just adds statistical power to the S900 based findings. Depending on the neurocognitive factor, the heritability analysis in the combined sample was conducted on 59-135 pairs of monozygotic twins (median=108.5) and 85-259 pairs of dizygotic twins and siblings (median=179). (A) Heritability of each trait. BMI has multiple estimates since it was entered into a bivariate analysis with each PPS separately. (B) Genetic correlations between BMI and each PPS. The genetic correlations are positive, because the PPS-s are designed to positively predict BMI. (C) Heritability of the significant phenotypic correlation between BMI and PPS. Horizontal lines depict 95% confidence intervals. Cogn=PPS of cognitive tests; corr=correlation; CT=PPS of cortical thickness; MTL=PPS of medial temporal lobe volume; Pers=PPS of personality tests; SC=PPS of subcortical structure volumes.

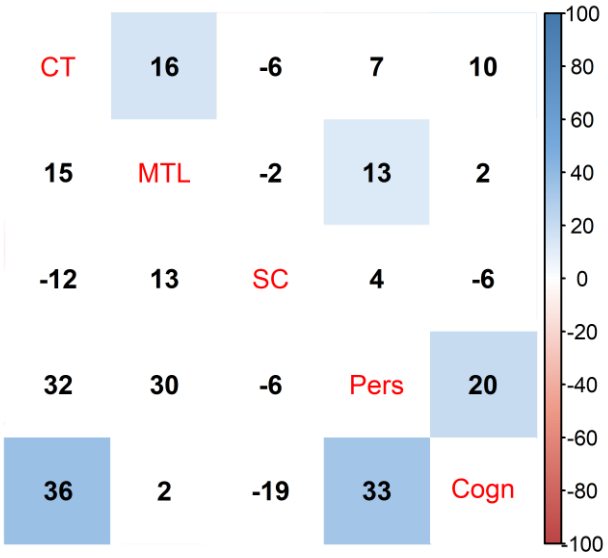

Fig. S11. Phenotypic (upper triangle) and genetic (lower triangle) correlations between PPS-s used for heritability analysis. Phenotypic correlations account for family structure. FDR-corrected significant correlations are highlighted with color. Correlations are multiplied by 100 for clarity. Cogn=PPS of cognitive tests; corr=correlation; CT=PPS of cortical thickness; MTL=PPS of medial temporal lobe volume; Pers=PPS of personality tests; SC=PPS of subcortical structure volumes

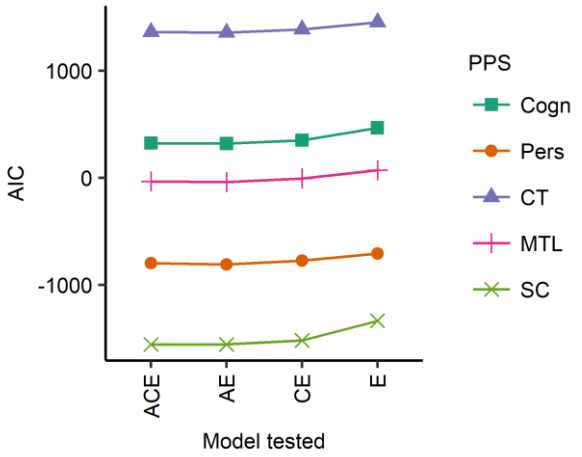

Fig. S12. Akaike Information Criteria (AIC) for BMI-PPS (poly-phenotype score) bivariate heritability decompositions. Cogn=PPS of cognitive tests; corr=correlation; CT=PPS of cortical thickness; MTL=PPS of medial temporal lobe volume; Pers=PPS of personality tests; SC=PPS of subcortical structure volumes.

Table S1. Descriptive statistics of samples analyzed.

Variable	S900	S1200n	S1200n unrelated
N	895	225	124
Age (years)	$\bar{x}$=28.83 (SD=3.67)	$\bar{x}$=28.85 (SD=3.84)	$\bar{x}$=29.31 (SD=3.83)
BMI (kg/m^2)	$\bar{x}$=27.27 (SD=5.77)	$\bar{x}$=26.51 (SD=5.21)	$\bar{x}$=26.32 (SD=5.18)
BMI groups			
Normal weight (BMI 18-24.9)	375 (41.9%)	101 (44.9%)	56 (45.2%)
Overweight (BMI 25-29.9)	285 (31.8%)	74 (32.9%)	45 (36.3%)
Obese (BMI 30+)	235 (26.3%)	50 (22.2%)	23 (18.5%)
Drug test positive			
No	777 (86.8%)	195 (86.7%)	105 (84.7%)
Yes	118 (13.2%)	30 (13.3%)	19 (15.3%)
Education (years)	$\bar{x}$=14.85 (SD=1.82)	$\bar{x}$=15.06 (SD=1.72)	$\bar{x}$=14.83 (SD=1.8)
Ethnicity:			
Hispanic/Latino	819 (91.5%)	198 (88%)	114 (91.9%)
Not Hispanic/Latino/unknown	76 (8.5%)	27 (12%)	10 (8.1%)
Families	384	151	66
1 sibling	37 (10.4%)	19 (20%)	19 (28.8%)
2 siblings	107 (30.1%)	49 (51.6%)	36 (54.5%)
3 siblings	163 (45.9%)	20 (21.1%)	11 (16.7%)
4 siblings	43 (12.1%)	6 (6.3%)	0 (0%)
5 siblings	5 (1.4%)	1 (1.1%)	0 (0%)
Gender			
Male	413 (46.1%)	120 (53.3%)	61 (49.2%)
Female no birth control	143 (16%)	24 (10.7%)	16 (12.9%)
Female with birth control	339 (37.9%)	81 (36%)	47 (37.9%)
Handedness	$\bar{x}$=65.07 (SD=45.13)	$\bar{x}$=68.93 (SD=41.03)	$\bar{x}$=70.73 (SD=36.97)
Income			
<$10,000	65 (7.3%)	16 (7.1%)	9 (7.3%)
10K-19,999	79 (8.8%)	12 (5.3%)	9 (7.3%)
20K-29,999	116 (13%)	24 (10.7%)	15 (12.1%)
30K-39,999	104 (11.6%)	30 (13.3%)	17 (13.7%)
40K-49,999	98 (10.9%)	23 (10.2%)	13 (10.5%)
50K-74,999	181 (20.2%)	46 (20.4%)	25 (20.2%)
75K-99,999	119 (13.3%)	28 (12.4%)	14 (11.3%)
>=100,000	133 (14.9%)	46 (20.4%)	22 (17.7%)
Race			

White	664 (74.2%)	176 (78.2%)	95 (76.6%)
Other/unknown	45 (5%)	21 (9.3%)	11 (8.9%)
Black or African Am.	145 (16.2%)	13 (5.8%)	8 (6.5%)
Asian/Nat. Hawaiian/Other Pacific Is.	41 (4.6%)	15 (6.7%)	10 (8.1%)

BMI=body mass index; Is=islander; Nat=native

Additional Dataset S1 (separate file)

See first tab of file "SI_Dataset_1.xlsx" for table of contents.

SI References

53. Barch DM, et al. (2013) Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage* 80:169–189.

54. Elam J (2017) HCP Data Dictionary Public. Available at: https://web.archive.org/web/20170425185833/https://wiki.humanconnectome.org/display/PublicData/HCP+Data+Dictionary+Public-+500+Subject+Release [Accessed April 25, 2017].

55. Ozer DJ, Benet-Martínez V (2006) Personality and the Prediction of Consequential Outcomes. *Annu Rev Psychol* 57(1):401–421.

56. BIC (2016) BIC - The McConnell Brain Imaging Centre: CIVET. *McConnell Brain Imaging Cent CIVET*. Available at: https://web.archive.org/web/20170505175011/http://www.bic.mni.mcgill.ca/ServicesSoftware/CIVET [Accessed December 21, 2016].

57. Zijdenbos AP, Forghani R, Evans AC (2002) Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE Trans Med Imaging* 21(10):1280–1291.

58. Sherif T, et al. (2014) CBRAIN: a web-based, distributed computing platform for collaborative neuroimaging research. *Front Neuroinformatics* 8. doi:10.3389/fninf.2014.00054.

59. Sled JG, Zijdenbos AP, Evans AC (1998) A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 17(1):87–97.

60. Tohka J, Zijdenbos A, Evans A (2004) Fast and robust parameter estimation for statistical partial volume models in brain MRI. *NeuroImage* 23(1):84–97.

61. Kim JS, et al. (2005) Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *NeuroImage* 27(1):210–221.

62. MacDonald D, Kabani N, Avis D, Evans AC (2000) Automated 3-D Extraction of Inner and Outer Surfaces of Cerebral Cortex from MRI. *NeuroImage* 12(3):340–356.

63. Lyttelton O, Boucher M, Robbins S, Evans A (2007) An unbiased iterative group registration template for cortical surface analysis. *NeuroImage* 34(4):1535–1544.

64. Lerch JP, Evans AC (2005) Cortical thickness analysis examined through power analysis and a population simulation. *NeuroImage* 24(1):163–173.

65. Klein A, Tourville J (2012) 101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol. *Front Neurosci* 6. doi:10.3389/fnins.2012.00171.

66. He Y, Chen ZJ, Evans AC (2007) Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. *Cereb Cortex N Y N 1991* 17(10):2407–2419.

67. Eyler LT, et al. (2012) A Comparison of Heritability Maps of Cortical Surface Area and Thickness and the Influence of Adjustment for Whole Brain Measures: A Magnetic Resonance Imaging Twin Study. *Twin Res Hum Genet Off J Int Soc Twin Stud* 15(3):304–314.

68. Coupe P, et al. (2008) An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans Med Imaging* 27(4):425–441.

69. Collins DL, Neelin P, Peters TM, Evans AC (1994) Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr* 18(2):192–205.

70. Weier K, Fonov V, Lavoie K, Doyon J, Collins DL (2014) Rapid automatic segmentation of the human cerebellum and its lobules (RASCAL)—Implementation and application of the patch-based label-fusion technique with a template library to segment the human cerebellum. *Hum Brain Mapp* 35(10):5026–5039.

71. Krishnan A, Williams LJ, McIntosh AR, Abdi H (2011) Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *NeuroImage* 56(2):455–475.

72. McIntosh AR, Bookstein FL, Haxby JV, Grady CL (1996) Spatial pattern analysis of functional brain images using partial least squares. *NeuroImage* 3(3 Pt 1):143–157.

73. Dudbridge F (2013) Power and predictive accuracy of polygenic risk scores. *PLoS Genet* 9(3):e1003348.

74. Ware EB, et al. (2017) Heterogeneity in polygenic scores for common human traits. *bioRxiv*:106062.

75. Dudbridge F, Newcombe PJ (2015) Accuracy of Gene Scores when Pruning Markers by Linkage Disequilibrium. *Hum Hered* 80(4):178–186.

76. R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/ [Accessed June 5, 2015].

77. Bates DM (2010) lme4: Mixed-effects modeling with R. *URL Httplme4 R-Forge R-Proj Orgbook*. Available at: http://www.wordlatex.com/static/templates/pdf/book_B.pdf [Accessed February 23, 2015].

78. Fox J, Weisberg HS (2010) *An R Companion to Applied Regression* (SAGE Publications, Inc, Thousand Oaks, Calif). Second Edition edition Available at:

http://socserv.socsci.mcmaster.ca/jfox/Books/Companion/.

79. Revelle W (2014) psych: Procedures for personality and psychological research. *Northwest Univ Evanst R Package Version* 1(1).

80. Wei T, Simko V (2016) *corrplot: Visualization of a Correlation Matrix* Available at: https://cran.r-project.org/web/packages/corrplot/index.html [Accessed January 25, 2017].

81. Wickham H (2009) *Ggplot2: elegant graphics for data analysis* (Springer, New York).

82. Wickham H (2011) The Split-Apply-Combine Strategy for Data Analysis. *J Stat Softw* 40(1). doi:10.18637/jss.v040.i01.

83. Wickham H, RStudio (2017) *tidyr: Easily Tidy Data with "spread()" and "gather()" Functions* Available at: https://cran.r-project.org/web/packages/tidyr/index.html [Accessed May 5, 2017].

84. Wilke CO, Wickham H (2016) *cowplot: Streamlined Plot Theme and Plot Annotations for "ggplot2"* Available at: https://cran.r-project.org/web/packages/cowplot/index.html [Accessed May 5, 2017].

85. Bartoń K (2016) *MuMIn: Multi-Model Inference* Available at: https://cran.r-project.org/web/packages/MuMIn/index.html.

86. Halekoh U, Højsgaard S (2014) A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models–the R package pbkrtest. *J Stat Softw* 59(9):1–32.

87. Heiberger TP and R (2016) *abind: Combine Multidimensional Arrays* Available at: https://CRAN.R-project.org/package=abind [Accessed March 26, 2018].

88. Wing MKC from J, et al. (2017) *caret: Classification and Regression Training* Available at: https://CRAN.R-project.org/package=caret [Accessed March 26, 2018].

89. Nowok B, Raab GM, Dibben C (2016) synthpop : Bespoke Creation of Synthetic Data in R. *J Stat Softw* 74(11). doi:10.18637/jss.v074.i11.

90. Phillips N (2017) *yarrr: A Companion to the e-Book "YaRrr!: The Pirate's Guide to R"* Available at: https://cran.r-project.org/web/packages/yarrr/index.html [Accessed February 13, 2017].

91. Therneau TM, Daniel S, Sinnwell J, Atkinson E (2015) *kinship2: Pedigree Functions* Available at: https://cran.r-project.org/web/packages/kinship2/index.html.

92. Schwartz M (2015) *WriteXLS: Cross-Platform Perl Based R Function to Create Excel 2003 (XLS) and Excel 2007 (XLSX) Files* Available at: https://CRAN.R-project.org/package=WriteXLS [Accessed March 28, 2018].

93. Worsley K, et al. (2009) Surfstat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random

field theory. NeuroImage 47 (Supplement 1), S102–S102, Organization for Human Brain Mapping. *2009 Annual Meeting*.

94. MATLAB (2016) (The MathWorks Inc., Natick, MA).

95. Neale MC, et al. (2016) OpenMx 2.0: Extended Structural Equation and Statistical Modeling. *Psychometrika* 81(2):535–549.

96. Institute for Behavioral Genetics (2017) International Behavioral Genetics Workshop. Available at: https://web.archive.org/web/20170426173719/http://www.colorado.edu/ibg/international-behavioral-genetics-workshop [Accessed March 28, 2017].