

Inter-individual differences in human brain structure and morphometry link to population variation in demographics and behavior

A.Llera^{1,2,3}, T. Wolfers^{1,2,3}, P. Mulders^{1,2,3}, C. F. Beckmann^{1,2,3,4}

1- Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands.

2- Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands.

3- Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen, The Netherlands.

4- Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), Oxford University

Abstract

We perform a comprehensive integrative analysis of multiple structural MR-based brain features and find strong evidence relating inter-individual structural variations to demographic and behavioral variates across a large cohort of healthy human volunteers. In particular, our findings shed some light on functional & structural integration, as we find a mode of structural variation that relates to and extends the ‘positive-negative’ behavioral spectrum which was recently reported as being associated with variations in functional connectivity.

Significance statement

This work provides for the first-time strong evidence relating human brain structure variations to a *wide range* of demographic and behavioral measures. We show that several measures previously associated to variation in functional MRI connectivity are in fact already associated at the structural level, pointing towards structure-function integration.

Introduction

Understanding individual human behavior has attracted the attention of scientists and philosophers since antiquity. The first quantitative approach intended to deepen such understanding dates to the first half of the 19-th century when ‘phrenology’^{1,2} related human behavior or cognitive abilities to skull measures. Technical, intellectual and clinical advances in the last two centuries allow us to now accurately quantify brain structure and function³⁻⁷, and to summarize certain ‘aspects’ of human behavior by means of standardized tests. Such advances facilitate automatized exploratory statistical learning analyses to uncover previously hidden relationships between brain features and human behavior, demographics or pathologies⁸. These developments are expected to be pushed even further with the emergence of the big data MR imaging epidemiology phenomenon^{9,10}, and some exponents of such expectations have already reported associations with blood-oxygen-level dependent (BOLD) brain function^{11,12}; for example, functional connectivity patterns can be used to identify individuals¹¹, predict fluid intelligence¹¹, or describe a mode of functional connectivity variation that relates to lifestyle, happiness and well-being¹². Although the brain’s function-structure relationships are not yet fully understood, linking structure to behavior is essential for either type of imaging modality to be fully interpretable as an imaging phenotype. Further, given the long-term character of some demographic variables (e.g. overall happiness), we hypothesize that different brain structural features such as regional variation in the density of gray matter or

subject-dependent degree of cortical expansion should also, to some extent, reflect these relationships with behavior. To test these hypotheses, in this work we make use of the large quantity of high quality behavioral and neuroimaging data collected by one of the big data initiatives, the Human Connectome Project⁹ (HCP). The HCP sample includes detailed structural imaging, diffusion MRI, resting-state and several different fMRI tasks for each subject. Furthermore, the availability of more than 300 behavioral and demographic measures allows the post-hoc exploration of a wide range of associations¹³. We further hypothesize that behavioral variations can be explained by more general brain structure variations than isolated single feature variations (e.g. cortical thickness variations); we consequently extract multiple structural features from the different MR modalities and perform a simultaneous analysis by linked independent component analysis (Linked ICA^{13,14}). Such analyses increase statistical power by evidence integration across different features^{15,16} and have been shown to be powerful in identifying correlated patterns of structural and diffusion spatial variation that can then be studied in relation to individual behavioral and demographic measures^{17,18,16}. Although similar analyses have been previously performed¹⁷, in this work we benefit from the unique characteristics of the data sample; we consider brain and behavioral data from close to 500 “healthy young adults” which reduces common pathology- and age-related variance and so increases the power to detect associations due to normal cross-sectional variability.

Our results support the hypothesis that structural brain features are strongly associated with demographic and behavioral variates. Interestingly, the most relevant mode of structural variation identified through the multi-modal data fusion approach relates to recent findings obtained using functional MRI data from the same HCP cohort. In particular, our findings closely resemble the ‘positive-negative’ set of behavioral measures identified in Smith et al.¹² on the basis of functional (co-)variations, pointing towards structure-function integration. Indeed, they support the notion that it is the spatial configuration in functional parcels rather than cross-parcel connectivity¹⁹ that is driving the brain-behavior relationships.

Methods

In this work, we use data from the Human Connectome Project (HCP) N=500 release which contains data from healthy young adults including twins and their non-twin siblings. In addition to performing more than 300 behavioral/demographic tests, each subject participated in structural, diffusion and several functional MRI recordings^{20,21}. A description of all MRI and behavioral/demographic measures included in our analysis can be found in van Essen et al.²⁰ and we provide a summary of the latter in the supplemental information (SI), Table S1. Due to structure-function integration we hypothesize that different biological features such as regional variation in the density of gray matter, white matter connectivity or subject dependent degree of cortical expansion should reflect similar associations with behavior as the ones reported at^{11,12}. To investigate such hypothesis, the structural MRI T1-weighted images were used to extract gray matter densities and cortical measures, using a Voxel Based Morphometry (VBM)^{22,23} (<http://www.fil.ion.ucl.ac.uk/spm12>) pipeline to extract cortical gray matter probability maps as well as maps of cortical thickness (CT) and pial area (PA)^{24,25} estimates by means of transforming all anatomical T1-weighted cortical surfaces through FreeSurfer v5.3 (<http://surfer.nmr.mgh.harvard.edu>). Further, the diffusion-weighted MRI data were used to extract several features, i.e. fractional anisotropy (FA), anisotropy mode (MO) and mean diffusivity (MD)^{26,27} (<https://fsl.fmrib.ox.ac.uk/fsl/v5.0.9>). In addition to these structural readouts

and in order to also include purely local morphometric differences across subjects, we also consider the images containing the Jacobian determinants (JD) of the warp fields defining the transformations of each subject's structural image onto a reference brain. These feature extraction operations are schematically summarized in Figure 1 operation A, and full details on the data processing performed to achieve each feature are provided in the SI under the section '*Individual features pre-processing*'. From the initial N=500 participants, several subjects were excluded on the basis of abnormalities in any of the features. In total, N=448 subjects were entered into further analyses. We then use the Linked-ICA model¹³ to simultaneously factorize the considered N=448 subjects' VBM, FA, MO, MD, CT, PA and JD features into independent sources (or components) of spatial variation. In brief, Linked-ICA is an extension of Bayesian ICA²⁸ to multiple input sets, where all individual ICA factorizations are linked through a shared common mixing matrix that reflect the subject-wise contribution to each component. This operation is represented in Figure 1 operation B where we can also appreciate that such factorization provides - per component - a set of spatial maps (one per feature modality), a vector of feature loadings that describe the degree to which the component is 'driven' by the different modalities, and a vector that describes how each individual subject contributes to a given component. Importantly, the subject-loadings define the cross-subject variation of the multi-modal effects and can subsequently be used to study relationships to other behavioral or demographic cross-subject variations by means of simple correlations. Given our sample size and following^{13,14} we report full results from a 100 dimensional factorization. Different model order decompositions were also performed to demonstrate the robustness to the choice of dimensionality; further, we also performed an analogous multi-modal analyses where we excluded the JD features as well as an independent component analyses²⁹ of the JD features in isolation to evaluate the dependency of the results on purely morphometric differences- these are reported in the SI. To perform the multi-modal analyses we developed a new implementation of the Linked ICA algorithm¹³ which will be made publicly available as part of the next release of the FSL toolbox²⁷.

Statistical analysis:

To uncover relationships between the behavioral/demographic measures and the components obtained from the Linked-ICA decomposition we perform a correlation analysis between each independent component subjects' contribution and each available behavioral measure. This operation is schematically summarized in Figure 1 operation C. To take into account the family structure present in the HCP sample while assessing significance we use the Permutation Analysis of Linear Models (PALM)^{30,31} and use 10^6 permutations per tested correlation (Figure 1 operation D). We define significance at $p < 0.05$ and address the multiple comparison by applying FDR correction³² as well as full Bonferroni correction (Figure 1 operation E).

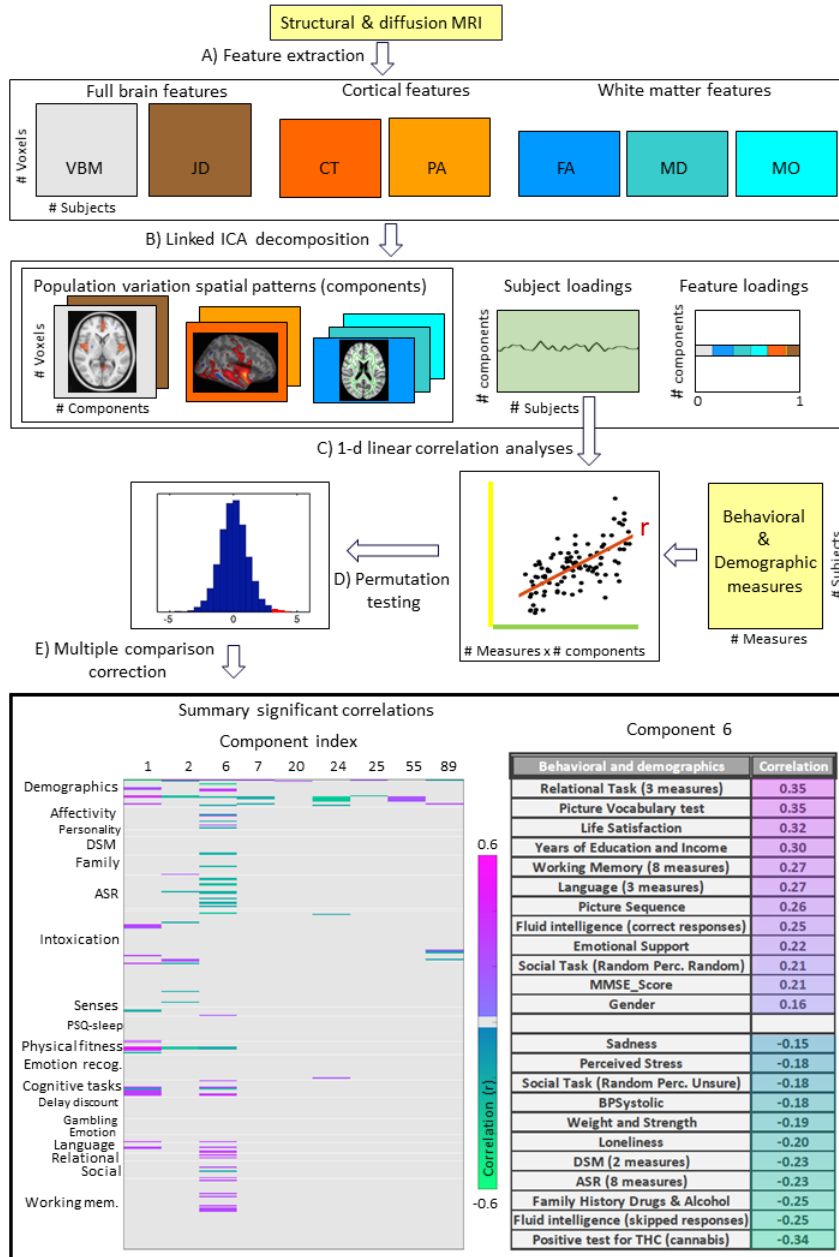


Figure 1: Data processing pipeline and main results. (A) Structural and diffusion-weighted MRI data are used to extract relevant features i.e. Voxel-Based Morphometry (VBM), Fractional Anisotropy (FA), Mean Diffusivity (MD), Anisotropy Mode (MO), Cortical Thickness (CT), Pial Area (PA) and Jacobian Determinants (JD). (B) These features are used as input to the Linked ICA algorithm. (C) Subject loadings of each independent component are fed together with the behavioral/demographic measures into a correlation analysis. The bottom left panel presents demographic and behavioral measures grouped by categories (y-axis), and a representative set of components reflecting significant correlation with at least one behavioral measure (x-axis). The color scale encodes the Pearson correlation coefficient and only significant correlations are color coded. In the bottom right panel we present a summary of component number 6 significant correlations to behavioral and demographic variates, resampling a mode of structural variation that links to and extends the ‘positive-negative’ behavioral spectrum previously attributed to functional connectivity variations¹².

Results

The multimodal brain data analyses resulted in a total of 100 collections of component maps, each of which can be represented as 7 spatial maps covering the gray-matter space (VBM), diffusion skeleton space (DTI, MD, MO), cortical vertex space (CT and PA) and a voxel-wise map of the Jacobian deformation (JD). In addition, each collection of maps is associated with a single vector of contributions that describe the degree to which a given collection is ‘driven’ by the different modalities. Further, each collection is associated with a single vector that describes how each individual subject contributes to the component. Post-hoc linear correlation analyses of these subject contributions with behavioral measures identified, after FDR correction¹, a total of 155 significant brain-behavior correlations, summarized by 30 components reflecting at least one significant relationship to behavior. We provide the full results in SI Table S2 and a brief summary in the bottom left panel of Figure 1 where we color code the significant Pearson correlation values for the components showing at least one Bonferroni corrected² significant correlation to a behavioral or demographic measure.

A single component, number 6, shows strong associations with 48 measures across several behavioral domains and across all structural modalities (SI Figure S2). In Figure 1 bottom right panel we provide a summary of the behavioral measures significantly correlating with component 6 as well as the corresponding Pearson correlation values. For interpretation, the behavioral measures are grouped and ordered according to a decreasing correlation value. Note that in the cases where several measures are grouped together we report their mean correlation value and full results can be again found in SI Table S2. We observe that component number 6 relates to various behavioral scores including working memory, language function and general wellbeing (life satisfaction, social support). In Figure 2 we present the associated spatial maps: VBM measures are most heavily weighted in bilateral orbitofrontal cortex, temporal pole, lingual gyrus and the putamen (first row). Morphometric differences (JD features) load into temporal lobes, caudate and brainstem (second row), and white matter tracts do most heavily weigh onto the internal capsule, anterior thalamic radiation and the anterior corona radiata (3rd, 4th, and 5th rows). Cortical effects (6th and 7th rows) are largely associations with multi-modal association cortex that show effects whereas primary sensory cortices are not implicated. Note that the involvement of areas such as the putamen and lingual gyrus are relevant to explain the behavioral relationships found with working memory and word processing. Further, the involvement of structural connections between subcortical and prefrontal areas as well as the orbitofrontal cortex and temporal poles could explain the link to more complex functions such emotional support or life satisfaction.

Components 1 and 2 also show to be related to several measures: components 1 relates mainly to gender, physical strength and language and it is defined by significant changes in gray matter density (VBM measures) and cortical real expansion (PA measure). Similarly, component 2 is driven by VBM maps and correlates with variations in gender, age, height, weight and strength. Components 7, 24, 25, 29 and 55 are driven by at least 3 feature modalities and they map into gender, weight, body mass and height. Number 89 maps VBM and JD into hematocrit and component 20 maps VBM and JD into age. Details about the spatial extent of each of these components can be found in SI, section ‘*spatial maps*’. Another set of components show

¹ FDR corrected, $q < 2.2 \times 10^{-4}$.

² Bonferroni corrected $q < 1.4 \times 10^{-6}$.

associations, albeit either being associated with only a single modality and/or with a small set of behavioral variates (SI Table S2). Many of these components show simple relationships to overall size measures, such as weight, body mass (BMI) or height, and the associations are weaker than the above reported; we consequently decided to not further discuss their spatial extent in this work and we provide full nifty images in the SI.

To validate the robustness of the presented results to the model order choice we performed analyses at different dimensionalities and observe that especially lower indexed components are highly reproducible. In particular, component number 6 is recovered at dimensionalities 90 and 110 with a subject mode correlation value of around $r \sim 0.9$ (details can be found in the SI section '*Robustness: model order*'). Regarding the influence of purely morphometric differences in the analyses, a comparative analysis excluding the JC revealed essentially unaltered brain-behavior associations. Further, an analyses of only the JD feature showed that no fully corrected significant association to the reported positive-negative structural mode is found when considering uniquely morphometric differences, even if considering several components together. However, uncorrected statistics suggests that information of the positive-negative mode could already be present at the morphometric level. These results are presented in SI section '*Robustness: analyses without Jacobians*', and '*Robustness: Analyzing morphometric differences*'.

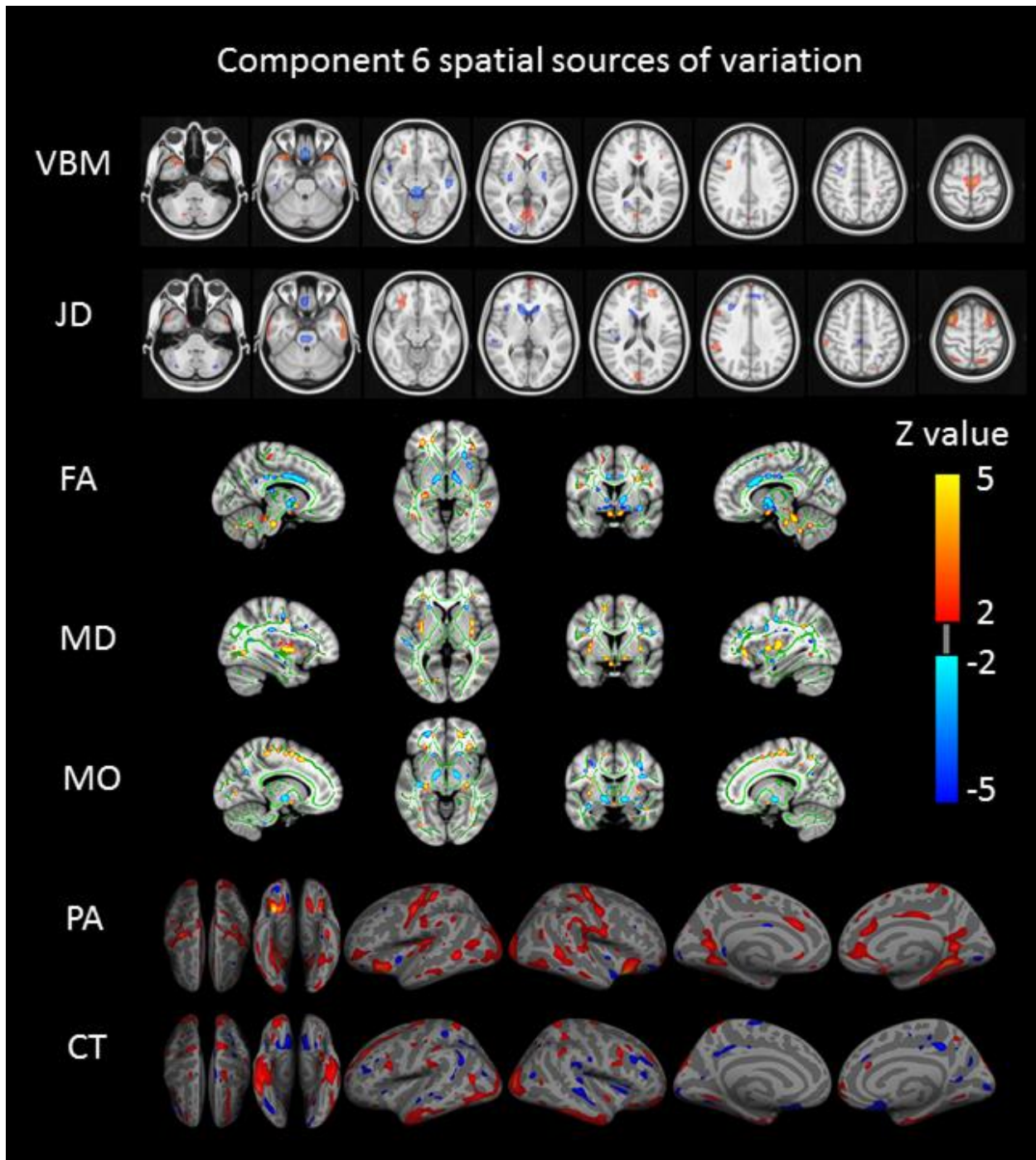


Figure 2: Component number 6 feature sources of variation. From top to bottom we visualize the VBM (Voxel Based Morphometry), JD (Jacobian Determinants), FA (Fractional Anisotropy), MD (Mean Diffusivity), MO (Mode of Anisotropy), PA (Pial Area), and CT (Cortical Thickness) spatial maps. For improved visualization, each modality has been thresholded at a z-value of 2. This mode of structural variation, component 6, that strongly reflects a ‘positive-negative’ behavioral spectrum, links to a wide range of brain regions across structural modalities and might reflect the structural multimodal foundation of a functional brain network linked to these variations that has been earlier identified.

Discussion

We present a simultaneous analysis of brain structural measures that reveals how several types of behavior and demographics link to variations in such measures of brain structure. Several components detect simple associations between brain size (encoded in gray matter density and cortical area) being related to gender, strength, endurance or language function.

More interestingly, we encounter a single pattern of gray and white matter covariation that is strongly associated with several measures relating to cognitive function including working memory and language function, while also being strongly related to several measures of wellbeing including life satisfaction or emotional support. Accordingly, the spatial organization of the component that relates to these measures predominantly includes regions and connections that are relevant to working memory and word processing such as the putamen and lingual gyrus^{33,34}. Additionally, the inclusion of regions such as the orbitofrontal cortex and temporal poles, as well as structural connections from subcortical to prefrontal regions, could explain the link to more complex functions such as emotional support and life satisfaction. Furthermore, the mode of structural variation we report here relates to several recently reported results obtained using functional MRI. In particular, our results relate to the ones presented in Finn et al.¹¹ since it identifies fluid intelligence measures and it also shares many behavioral measures also identified by the ‘positive-negative’ mode reported in Smith et al.¹². Clearly, the functional analyses presented in Smith et al.¹² and the one we just presented here, while using entirely different MRI measurements, are both able to get at the core of the same behavioral spectrum; our analyses reliably augments the spectrum of behavioral variables reported by the functional analyses by extending it with many working memory, language, relational task, ASR and DSM measures (Figure 1 bottom right and SI Table S2). It is to note here that while the statistics reported in Smith et al.¹² were obtained from a Canonical Correlation Analyses (CCA) between partial correlation matrices and all behavioral measures at once, the statistics we present here involve simple linear correlations. While the former type of analysis can benefit from the multi-variate type of analysis through the application of CCA, ensuing results can be hard to interpret. The straight-forward individual linear correlation analysis against the behavioral/demographic measures separately instead affords simple interpretation, albeit possibly being over-conservative given the chosen significance level.

These findings directly look into the relationship between brain structure and function. In fact, the functional mode of variation is strongly associated with connectivity in brain areas approximately resembling the Default Mode Network and, given the spatial extent and the strong weight of the DWI data in the structural mode we report, it seems reasonable to assume that these white matter structure variations could contribute to the functional connectivity changes reported in Smith et al.¹². Further we found no clear spatial overlap between the reported structural mode and the cortical functional extent of the ‘positive-negative’ mode, suggesting that functional-structural simultaneous analyses should increase the sensitivity of both, functional and structural analyses. Further, these results might question whether group functional connectivity measures using fMRI provide direct measures of brain connectivity or are biased due to individual structural differences that have been taken into account inadequately. An analogous multimodal analysis excluding the JD feature provided equivalent results to the presented here (SI Table S3) and uni-modal analysis of uniquely the JD features (using simple ICA-based decomposition²⁹ of the single JD modality) did not provide significant correlation to the behavioral mode at the level of fully corrected

statistics. These extra analyses confirm that the structural features relating to the behavioral mode are not uniquely driven by morphometric differences and consequently suggest that the behavioral associations found using functional MRI, although could be influenced, are not uniquely driven by the structural differences; these results seem to be aligned with recent findings by Bijsterbosch et al.¹⁹ where it is shown that individual spatial configurations extracted from functional MRI rather than the connectivity profiles between areas seem to stronger relate to the positive-negative mode.

In a future study, we will couple structural-functional modalities to investigate the link of these modalities more immediately. Here, we showed that a positive-negative mode of behavioral variation extends beyond the functional domain to also link to multimodal brain structure. This will have important implications in the future analyses of neuroimaging big data and will help improve our understanding over the functional-structural integration.

Acknowledgements

We are grateful to Stephen M. Smith for the helpful discussions and to V. Kumar for help with the visualization of the DWI results. The research leading to these results has received funding through a Synergy Grant by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013), ERC Grant Agreement no. 319456. We further gratefully acknowledge support from the Netherlands Organization for Scientific Research (NWO) through VIDI grant to CFB (864.12.003) and we also gratefully acknowledge funding from the Wellcome Trust UK Strategic Award (098369/Z/12/Z).

References

1. Simpson, D. Phrenology and the neurosciences: Contributions of F. J. Gall and J. G. Spurzheim. in *ANZ Journal of Surgery* **75**, 475–482 (2005).
2. Fodor, J. A. *The modularity of mind: an essay on faculty psychology. Reasoning: Studies of human inference and its foundations* (1983). doi:10.2307/413815
3. Lerch, J. P. *et al.* Studying neuroanatomy using MRI. *Nat. Neurosci.* **20**, 314–326 (2017).
4. Scott A. Huettel; Allen W. Song; Gregory McCarthy *et al.* *Functional Magnetic Resonance Imaging. Book* **23**, (2004).
5. Friston, K. J. *et al.* Classical and Bayesian inference in neuroimaging: applications. *Neuroimage* **16**, 484–512 (2002).
6. Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M. & Smith, S. M. Multilevel linear modelling for fMRI group analysis using Bayesian inference. *Neuroimage* **21**, 1732–1747 (2004).
7. Rorden, C., Bonilha, L. & Nichols, T. E. Rank-order versus mean based

- statistics for neuroimaging. *Neuroimage* **35**, 1531–1537 (2007).
8. Poldrack, R. A. & Farah, M. J. Progress and challenges in probing the human brain. *Nature* **526**, 371–379 (2015).
 9. Van Essen, D. C. *et al.* The WU-Minn Human Connectome Project: An overview. *Neuroimage* **80**, 62–79 (2013).
 10. Collins, R. What makes UK Biobank special? *The Lancet* **379**, 1173–1174 (2012).
 11. Finn, E. S. *et al.* Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* **18**, 1664–1671 (2015).
 12. Smith, S. M. *et al.* A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat. Neurosci.* **18**, 1–7 (2015).
 13. Groves, A. R., Beckmann, C. F., Smith, S. M. & Woolrich, M. W. Linked independent component analysis for multimodal data fusion. *Neuroimage* **54**, 2198–2217 (2011).
 14. Groves, A. R. *et al.* Benefits of multi-modal fusion analysis on a large-scale dataset: Life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure. *Neuroimage* **63**, 365–380 (2012).
 15. Wolfers, T. *et al.* Refinement by integration : aggregated effects of multimodal imaging markers on adult ADHD. 1–9 (2017).
doi:10.1503/jpn.160240
 16. Doan, N. T. *et al.* Dissociable diffusion MRI patterns of white matter microstructure and connectivity in Alzheimer’s disease spectrum. *Sci. Rep.* **7**, 45131 (2017).
 17. Douaud, G. *et al.* A common brain network links development, aging, and vulnerability to disease. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 17648–17653 (2014).
 18. Francx, W. *et al.* Integrated analysis of gray and white matter alterations in attention-deficit/hyperactivity disorder. *NeuroImage Clin.* **11**, (2016).
 19. Bijsterbosch, J. D. *et al.* The relationship between spatial configuration and functional connectivity of brain regions. *Elife* **7**, (2018).

20. Van Essen, D. C. *et al.* The Human Connectome Project: A data acquisition perspective. *NeuroImage* **62**, 2222–2231 (2012).
21. Elam, J. S. & Van Essen, P. D. Human Connectome Project. in *Encyclopedia of Computational Neuroscience* 1–4 (2013).
22. Ashburner, J. & Friston, K. J. Voxel-Based Morphometry—The Methods. *Neuroimage* **11**, 805–821 (2000).
23. Ashburner, J. & Friston, K. J. Unified segmentation. *Neuroimage* **26**, 839–851 (2005).
24. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *Neuroimage* **9**, 179–194 (1999).
25. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System. *Neuroimage* **9**, 195–207 (1999).
26. Smith, S. M. *et al.* Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage* **31**, 1487–1505 (2006).
27. Jenkinson, M., Beckmann, C., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–90 (2012).
28. Choudrey, R. Variational Methods for Bayesian Independent Component Analysis. (University of Oxford).
29. Beckmann, C. F. & Smith, S. M. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging* **23(2)**, 137–152 (2004).
30. Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. Permutation inference for the general linear model. *Neuroimage* **92**, 381–397 (2014).
31. Winkler, A. M., Webster, M. A., Vidaurre, D., Nichols, T. E. & Smith, S. M. Multi-level block permutation. *Neuroimage* **123**, 253–268 (2015).
32. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* **57**, 289–300 (1995).
33. Arsalidou, M., Duerden, E. G. & Taylor, M. J. The centre of the brain:

- Topographical model of motor, cognitive, affective, and somatosensory functions of the basal ganglia. *Hum. Brain Mapp.* **34**, 3031–3054 (2013).
34. Mechelli, A., Humphreys, G. W., Mayall, K., Olson, A. & Price, C. J. Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. *Proc Biol Sci* **267**, 1909–1913 (2000).
 35. Ashburner, J. A fast diffeomorphic image registration algorithm. *Neuroimage* **38**, 95–113 (2007).
 36. Fischl, B. & Dale, A. M. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc. Natl. Acad. Sci.* **97**, 11050–11055 (2000).
 37. Fischl, B. *et al.* Cortical folding patterns and predicting cytoarchitecture. *Cereb. Cortex* **18**, 1973–1980 (2008).
 38. Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23**, S208–19 (2004).

Supplemental information (SI)

In Table S1 we provide a summary of the behavioral and demographic measures present in the Human Connectome Project (HCP) sample. For easier interpretation we grouped them here by categories and a full detailed description can be found in van Essen et al.²⁰.

Social & Demographics	Gender	Intoxication	Total_Hard_Liquor_7days	Reading & vocabulary	ReadEng_Unadj
	Age		Avg_Weekday_Hard_Liquor_7days		ReadEng_AgeAdj
	HasGT		Avg_Weekend_Hard_Liquor_7days		PicVocab_Unadj
	Race		Total_Other_Alc_7days		PicVocab_AgeAdj
	Ethnicity		Avg_Weekday_Other_Alc_7days	Processing speed	ProcSpeed_Unadj
	Handedness		Avg_Weekend_Other_Alc_7days		ProcSpeed_AgeAdj
	SSAGA_Employ		SSAGA_Alc_D4_Dp_Sx	Delay discount	DDisc_SV_1mo_200
	SSAGA_Income		SSAGA_Alc_D4_Ab_Dx		DDisc_SV_6mo_200
	SSAGA_Educ		SSAGA_Alc_D4_Ab_Sx		DDisc_SV_1yr_200
	SSAGA_InSchool		SSAGA_Alc_D4_Dp_Dx		DDisc_SV_3yr_200
	SSAGA_Rlshp		SSAGA_Alc_12_Drinks_Per_Day		DDisc_SV_5yr_200
	SSAGA_MOBorn		SSAGA_Alc_12_Frq		DDisc_SV_10yr_200
	Height		SSAGA_Alc_12_Frq_5plus		DDisc_SV_1mo_40K
	Weight		SSAGA_Alc_12_Frq_Drk		DDisc_SV_6mo_40K
	BMI		SSAGA_Alc_12_Max_Drinks		DDisc_SV_1yr_40K
	SSAGA_BMICat		SSAGA_Alc_Age_1st_Use		DDisc_SV_3yr_40K
	SSAGA_BMICatHeaviest		SSAGA_Alc_Hvy_Drinks_Per_Day	DDisc_SV_5yr_40K	
	Blood_Drawn		SSAGA_Alc_Hvy_Frq	DDisc_SV_10yr_40K	
	Hematocrit_1		SSAGA_Alc_Hvy_Frq_5plus	DDisc_AUC_200	
	BPSystolic		SSAGA_Alc_Hvy_Frq_Drk	DDisc_AUC_40K	
BPDiastolic	SSAGA_Alc_Hvy_Max_Drinks	Gambling	Gambling_Task_Perc_Larger		
AngAffect_Unadj	Total_Any_Tobacco_7days		Gambling_Task_Perc_Smaller		
AngHostil_Unadj	Times_Used_Any_Tobacco_Today		Gambling_Task_Median_RT_Larger		
AngAggr_Unadj	Num_Days_Used_Any_Tobacco_7days		Gambling_Task_Median_RT_Smaller		
FearAffect_Unadj	Avg_Weekday_Any_Tobacco_7days		Gambling_Task_Reward_Perc_Larger		
FearSomat_Unadj	Avg_Weekend_Any_Tobacco_7days		Gambling_Task_Reward_Median_RT_Larger		
Sadness_Unadj	Total_Cigarettes_7days		Gambling_Task_Reward_Perc_Smaller		
LifeSatisf_Unadj	Avg_Weekday_Cigarettes_7days		Gambling_Task_Reward_Median_RT_Smaller		

Affective	MeanPurp_Unadj	Intoxication	Avg_Weekend_Cigarettes_7days	Gambling	Gambling_Task_Punish_Perc_Larger
	PosAffect_Unadj		Total_Cigars_7days		Gambling_Task_Punish_Median_RT_Larger
	Friendship_Unadj		Avg_Weekday_Cigars_7days		Gambling_Task_Punish_Perc_Smaller
	Loneliness_Unadj		Avg_Weekend_Cigars_7days		Gambling_Task_Punish_Median_RT_Smaller
	PercHostil_Unadj		Total_Chew_7days	Emotion	Emotion_Task_Acc
	PercReject_Unadj		Avg_Weekday_Chew_7days		Emotion_Task_Median_RT
	EmotSupp_Unadj		Avg_Weekend_Chew_7days		Emotion_Task_Face_Acc
	InstruSupp_Unadj		Total_Snuff_7days		Emotion_Task_Face_Median_RT
	PercStress_Unadj		Avg_Weekday_Snuff_7days		Emotion_Task_Shape_Acc
	SelfEff_Unadj		Avg_Weekend_Snuff_7days		Emotion_Task_Shape_Median_RT
Personality	NEOFAC_A	Total_Other_Tobacco_7days	Language	Language_Task_Acc	
	NEOFAC_O	Avg_Weekday_Other_Tobacco_7days		Language_Task_Median_RT	
	NEOFAC_C	Avg_Weekend_Other_Tobacco_7days		Language_Task_Story_Acc	
	NEOFAC_N	SSAGA_TB_Smoking_History		Language_Task_Story_Median_RT	
	NEOFAC_E	SSAGA_TB_Still_Smoking		Language_Task_Story_Avg_Difficulty_Level	
DSM	DSM_Depr_Raw	SSAGA_Times_Used_Illicits	Relational	Language_Task_Math_Acc	
	DSM_Depr_Pct	SSAGA_Times_Used_Cocaine		Language_Task_Math_Median_RT	
	DSM_Anxi_Raw	SSAGA_Times_Used_Hallucinogens		Language_Task_Math_Avg_Difficulty_Level	
	DSM_Anxi_Pct	SSAGA_Times_Used_Opiates	Social	Relational_Task_Acc	
	DSM_Somp_Raw	SSAGA_Times_Used_Sedatives		Relational_Task_Median_RT	
	DSM_Somp_Pct	SSAGA_Times_Used_Stimulants		Relational_Task_Match_Acc	
	DSM_Avoid_Raw	SSAGA_Mj_Use		Relational_Task_Match_Median_RT	
	DSM_Avoid_Pct	SSAGA_Mj_Ab_Dep		Relational_Task_Rel_Acc	
	DSM_Adh_Raw	SSAGA_Mj_Times_Used		Relational_Task_Rel_Median_RT	
	DSM_Adh_Pct	Senses		Noise_Comp	Social_Task_Perc_Random
DSM_Inat_Raw	Odor_Unadj		Social_Task_Perc_TOM		
DSM_Hype_Raw	Odor_AgeAdj		Social_Task_Perc_Unsure		
DSM_Antis_Raw	PainInterf_Tscore		Social_Task_Median_RT_Random		
DSM_Antis_Pct	Taste_Unadj	Social_Task_Median_RT_TOM			
family history	FamHist_Moth_Scz	Taste_AgeAdj	Social_Task_Random_Perc_Random		
	FamHist_Moth_Dep	Visual acuity	Social_Task_Random_Median_RT_Random		
	FamHist_Fath_Dep		Social_Task_Random_Perc_TOM		

family history	FamHist_Moth_BP	MMSE	MMSE_Score	Social	Social_Task_Random_Perc_Unsure
	FamHist_Fath_BP		PSQI_Score		Social_Task_TOM_Perc_Random
	FamHist_Moth_AnX		PSQI_Min2Asleep		Social_Task_TOM_Perc_TOM
	FamHist_Fath_AnX		PSQI_AmtSleep		Social_Task_TOM_Median_RT_TOM
	FamHist_Moth_DrgAlc		PSQI_Latency30Min		Social_Task_TOM_Perc_Unsure
	FamHist_Fath_DrgAlc	Sleep			WM_Task_Acc
	FamHist_Moth_Alz			PSQI_WakeUp	WM_Task_Median_RT
	FamHist_Fath_Alz			PSQI_Bathroom	WM_Task_2bk_Acc
	FamHist_Moth_PD			PSQI_Snore	WM_Task_2bk_Median_RT
	FamHist_Fath_PD			PSQI_TooCold	WM_Task_0bk_Acc
	FamHist_Moth_None			PSQI_TooHot	WM_Task_0bk_Median_RT
	FamHist_Fath_None			PSQI_BadDream	WM_Task_0bk_Body_Acc
ASR_AnXd_Raw			PSQI_Pain	WM_Task_0bk_Body_Acc_Target	
ASR_AnXd_Pct			PSQI_Other	WM_Task_0bk_Body_Acc_Nontarget	
ASR_Witd_Raw			PSQI_Quality	WM_Task_0bk_Face_Acc	
ASR_Witd_Pct		PSQI_SleepMeds	WM_Task_0bk_Face_Acc_Target		
ASR_Soma_Raw		PSQI_DayStayAwake	WM_Task_0bk_Face_ACC_Nontarget		
ASR_Soma_Pct		PSQI_DayEnthusiasm	WM_Task_0bk_Place_Acc		
ASR_Thot_Raw		PSQI_BedPtnrRmate	WM_Task_0bk_Place_Acc_Target		
ASR_Thot_Pct	Physical fitness			WM_Task_0bk_Place_Acc_Nontarget	
ASR_Attn_Raw			Endurance_Unadj	WM_Task_0bk_Tool_Acc	
ASR_Attn_Pct			Endurance_AgeAdj	WM_Task_0bk_Tool_Acc_Target	
ASR_Aggr_Raw			GaitSpeed_Comp	WM_Task_0bk_Tool_Acc_Nontarget	
ASR_Aggr_Pct			Dexterity_Unadj	WM_Task_2bk_Body_Acc	
ASR_Rule_Raw			Dexterity_AgeAdj	WM_Task_2bk_Body_Acc_Target	
ASR_Rule_Pct		Strength_Unadj	WM_Task_2bk_Body_Acc_Nontarget		
ASR_Intr_Raw	Spatial orientation		VSPLOT_TC	WM_Task_2bk_Face_Acc	
ASR_Intr_Pct			VSPLOT_CRTE	WM_Task_2bk_Face_Acc_Target	
ASR_Oth_Raw			VSPLOT_OFF	WM_Task_2bk_Face_Acc_Nontarget	
ASR_Crit_Raw	Sustained attention		SCPT_TP	WM_Task_2bk_Place_Acc	
ASR_Intn_Raw			SCPT_TN	WM_Task_2bk_Place_Acc_Target	
ASR_Intn_T			SCPT_FP	WM_Task_2bk_Place_Acc_Nontarget	
ASR_Extn_Raw			SCPT_FN	WM_Task_2bk_Tool_Acc	
ASR_Extn_T			SCPT_TPRT	WM_Task_2bk_Tool_Acc_Target	
ASR_TAO_Sum			SCPT_SEN	WM_Task_2bk_Tool_Acc_Nontarget	
ASR_Totp_Raw			SCPT_SPEC	WM_Task_0bk_Body_Median_RT	

ASR	ASR_Totp_T	Sustained attention	SCPT_LRNR		
SSGA - other	SSAGA_ChildhoodConduct	verball episodic memory	IWRD_TOT	WM_Task_Obk_Body_Median_RT_Target	
	SSAGA_PanicDisorder		IWRD_RTC	WM_Task_Obk_Body_Median_RT_Nontarget	
Intoxication	Cocaine	List sorting	ListSort_Unadj	WM_Task_Obk_Face_Median_RT	
	THC		ListSort_AgeAdj	WM_Task_Obk_Face_Median_RT_Target	
	Opiates	Emotion recognition	ER40_CR	WM_Task_Obk_Face_Median_RT_Nontarget	
	Amphetamines		ER40_CR	WM_Task_Obk_Face_Median_RT_Nontarget	
	Methamphetamine		ER40ANG	WM_Task_Obk_Face_Median_RT_Nontarget	
	Oxycontin		ER40FEAR	WM_Task_Obk_Face_Median_RT_Nontarget	
	Total_Drinks_7days		ER40HAP	WM_Task_Obk_Face_Median_RT_Nontarget	
	Num_Days_Drank_7days		ER40NOE	WM_Task_Obk_Face_Median_RT_Nontarget	
	Avg_Weekday_Drinks_7days		ER40SAD	WM_Task_Obk_Face_Median_RT_Nontarget	
	Avg_Weekend_Drinks_7days		Picture sequence	PicSeq_Unadj	WM_Task_Obk_Face_Median_RT_Nontarget
	Total_Beer_Wine_Cooler_7days			PicSeq_AgeAdj	WM_Task_Obk_Face_Median_RT_Nontarget
	Avg_Weekday_Beer_Wine_Cooler_7days		Card sorting	CardSort_Unadj	WM_Task_Obk_Face_Median_RT_Nontarget
	Avg_Weekend_Beer_Wine_Cooler_7days	CardSort_AgeAdj		WM_Task_Obk_Face_Median_RT_Nontarget	
	Total_Malt_Liquor_7days	Flanker	Flanker_Unadj	WM_Task_Obk_Face_Median_RT_Nontarget	
	Avg_Weekday_Malt_Liquor_7days		Flanker_AgeAdj	WM_Task_Obk_Face_Median_RT_Nontarget	
	Avg_Weekend_Malt_Liquor_7days	fluid intelligence	PMAT24_A_CR	WM_Task_Obk_Face_Median_RT_Nontarget	
	Total_Wine_7days			PMAT24_A_SI	WM_Task_Obk_Face_Median_RT_Nontarget
	Avg_Weekday_Wine_7days			PMAT24_A_RTCR	WM_Task_Obk_Face_Median_RT_Nontarget
	Avg_Weekend_Wine_7days				WM_Task_Obk_Face_Median_RT_Nontarget
					Working memory
				WM_Task_2bk_Body_Median_RT	
				WM_Task_2bk_Body_Median_RT_Target	
				WM_Task_2bk_Body_Median_RT_Nontarget	
				WM_Task_2bk_Face_Median_RT	
				WM_Task_2bk_Face_Median_RT_Target	
				WM_Task_2bk_Face_Median_RT_Nontarget	
				WM_Task_2bk_Place_Median_RT	
				WM_Task_2bk_Place_Median_RT_Target	
				WM_Task_2bk_Place_Median_RT_Nontarget	
				WM_Task_2bk_Tool_Median_RT	
				WM_Task_2bk_Tool_Median_RT_Target	
				WM_Task_2bk_Tool_Median_RT_Nontarget	

Table S1: Summary of the behavioral/demographic measures present in the HCP sample.

Individual features pre-processing:

Prior to gray matter volume estimation, all participants' T1 images were rigidly aligned using statistical parametric mapping software (SPM-12). Subsequently, images were segmented, normalized, and bias-field-corrected using 'new segment' from VBM-SPM12 (<http://www.fil.ion.ucl.ac.uk/spm>, ^{22,23}), yielding images containing gray and white matter segments plus CSF. DARTEL ³⁵ was then used to create a study-

specific gray matter template to which all segmented images were normalized. Subsequently, all gray matter volumes were smoothed with a 9.4 mm FWHM Gaussian smoothing kernel ($\sigma=4$ mm).

Structural MRI images were fed into FreeSurfer v5.3 software to extract measures for cortical thickness and areal expansion^{24,25} (<http://surfer.nmr.mgh.harvard.edu>). The standard FreeSurfer preprocessing pipeline (recon-all) was applied to these images, in which a reconstruction of the cortical sheet was estimated using intensity and continuity information. Cortical thickness was determined as the closest distance from the gray/white boundary to the gray/cerebrospinal fluid (CSF) boundary at each vertex³⁶. Surface area in FreeSurfer is estimated as relative amount of expansion or compression at each vertex when registering each participant's surface to a common atlas. Surface maps were resampled and mapped to a common coordinate system³⁷. During preprocessing, the data were registered onto the high-resolution average participant surface space (fsaverage), and a 10 mm FWHM surface-based smoothing kernel was applied.

Further, the Jacobian images for each subject are directly available from the HCP repository and the diffusion weighted data (DWI) was preprocessed using the DTIFIT routine from FSL^{27,38} (<https://fsl.fmrib.ox.ac.uk/fsl>) to create the FA, MO and MD images that were then feed into the TBSS pipeline²⁶.

Finally, for computational reasons^{13,14}, the VBM images were spatially down sampled to 4mm isotropic and the DTI images to 2mm isotropic voxels.

Main results

In Table S2 we summarize the significant results obtained (FDR corrected $q < 2.2 \times 10^{-4}$). From left to right columns we present the component number, behavioral/demographical measure, correlation value, and the permutation p-value (PALM).

Component	Behavior/demographics	correlation	P value
1	Gender	-0.69181	1,00E-06
1	SSAGA_Employ	0.18407	1,00E-05
1	SSAGA_Income	0.20703	1,00E-06
1	Height	0.62158	1,00E-06
1	Weight	0.3008	1,00E-06
1	Hematocrit_1	0.30331	1,00E-06
1	Total_Beer_Wine_Cooler_7days	0.26895	1,00E-06
1	Avg_Weekday_Beer_Wine_Cooler_7days	0.24248	1,00E-06
1	Avg_Weekend_Beer_Wine_Cooler_7days	0.22384	1,00E-05
1	SSAGA_Alc_12_Max_Drinks	0.26547	2,00E-05
1	SSAGA_Alc_Hvy_Max_Drinks	0.31291	1,00E-06
1	Taste_Unadj	-0.22246	3,00E-05
1	Taste_AgeAdj	-0.22173	5,00E-05

[Type here]

1	Endurance_Unadj	0.31194	1,00E-06
1	Endurance_AgeAdj	0.30377	1,00E-06
1	Strength_Unadj	0.52027	1,00E-06
1	Strength_AgeAdj	0.53415	1,00E-06
1	VSPLIT_TC	0.25689	1,00E-06
1	VSPLIT_OFF	-0.25987	1,00E-06
1	PMAT24_A_CR	0.22625	1,00E-05
1	PMAT24_A_SI	-0.20436	2,00E-05
1	PMAT24_A_RTICR	0.1833	0.00018
1	ReadEng_Unadj	0.30153	1,00E-06
1	ReadEng_AgeAdj	0.30711	1,00E-06
1	PicVocab_Unadj	0.27521	1,00E-06
1	PicVocab_AgeAdj	0.29513	1,00E-06
1	Language_Task_Acc	0.35065	1,00E-06
1	Language_Task_Story_Avg_Difficulty_Level	0.24281	1,00E-05
1	Language_Task_Math_Acc	0.3251	1,00E-06
2	Gender	0.36923	1,00E-06
2	Age	-0.30654	1,00E-06
2	Height	-0.30887	1,00E-06
2	Weight	-0.28027	1,00E-06
2	FamHist_Fath_None	0.19331	0.00012
2	ASR_Rule_Raw	-0.19904	8,00E-05
2	Avg_Weekday_Drinks_7days	-0.18167	0.00016
2	SSAGA_Alc_Hvy_Frq	0.19232	0.00016
2	SSAGA_Alc_Hvy_Frq_5plus	0.22835	1,00E-05
2	SSAGA_Alc_Hvy_Max_Drinks	-0.2049	3,00E-05
2	SSAGA_TB_Smoking_History	-0.20967	8,00E-05
2	SSAGA_Mj_Use	-0.21131	2,00E-05
2	Strength_Unadj	-0.28766	1,00E-06
2	Strength_AgeAdj	-0.33854	1,00E-06
3	Weight	-0.19881	0.00012
3	Dexterity_Unadj	0.22919	7,00E-05
3	Dexterity_AgeAdj	0.2275	5,00E-05
3	Relational_Task_Acc	0.20415	0.00022
3	Relational_Task_Match_Acc	0.20791	0.00013
6	Gender	0.16981	3,00E-05
6	SSAGA_Income	0.2709	1,00E-06
6	SSAGA_Educ	0.34143	1,00E-06
6	Weight	-0.21022	8,00E-05

6	BPSystolic	-0.18174	0.0002
6	Sadness_Unadj	-0.14847	7,00E-05
6	LifeSatisf_Unadj	0.3216	1,00E-06
6	Loneliness_Unadj	-0.20117	4,00E-05
6	EmotSupp_Unadj	0.22221	0.00013
6	PercStress_Unadj	-0.18487	1,00E-05
6	DSM_Antis_Raw	-0.23362	1,00E-05
6	DSM_Antis_Pct	-0.2283	3,00E-05
6	FamHist_Fath_DrgAlc	-0.25365	0.00019
6	ASR_Witd_Raw	-0.24438	1,00E-05
6	ASR_Witd_Pct	-0.22928	1,00E-05
6	ASR_Thot_Raw	-0.29199	1,00E-05
6	ASR_Thot_Pct	-0.28879	1,00E-05
6	ASR_Rule_Raw	-0.2399	1,00E-06
6	ASR_Rule_Pct	-0.22106	2,00E-05
6	ASR_Crit_Raw	-0.20969	8,00E-05
6	ASR_Extn_Raw	-0.19428	2,00E-05
6	ASR_Extn_T	-0.18827	1,00E-05
6	ASR_Totp_Raw	-0.1787	0.00017
6	THC	-0.33929	1,00E-06
6	MMSE_Score	0.21061	2,00E-05
6	Strength_Unadj	-0.18765	8,00E-05
6	Strength_AgeAdj	-0.17876	0.00018
6	PicSeq_AgeAdj	0.25965	0.00012
6	PMAT24_A_CR	0.25373	0.00016
6	PMAT24_A_SI	-0.25185	0.00014
6	PicVocab_Unadj	0.35269	1,00E-06
6	PicVocab_AgeAdj	0.34436	1,00E-05
6	Language_Task_Acc	0.28331	3,00E-05
6	Language_Task_Story_Avg_Difficulty_Level	0.26322	8,00E-05
6	Language_Task_Math_Avg_Difficulty_Level	0.2667	1,00E-06
6	Relational_Task_Acc	0.37603	1,00E-06
6	Relational_Task_Match_Acc	0.29481	1,00E-05
6	Relational_Task_Rel_Acc	0.33171	1,00E-06
6	Social_Task_Random_Perc_Random	0.21119	7,00E-05
6	Social_Task_Random_Perc_Unsure	-0.18456	0.00021
6	WM_Task_Acc	0.29494	2,00E-05
6	WM_Task_2bk_Acc	0.29779	8,00E-05
6	WM_Task_0bk_Place_Acc	0.23687	1,00E-05

6	WM_Task_0bk_Place_Acc_Nontarget	0.26853	1,00E-05
6	WM_Task_2bk_Face_Acc	0.29316	1,00E-05
6	WM_Task_2bk_Face_Acc_Nontarget	0.24898	0.0001
6	WM_Task_2bk_Place_Acc	0.2858	1,00E-05
6	WM_Task_2bk_Place_Acc_Target	0.21687	0.0002
7	Gender	0.26174	1,00E-06
7	Weight	-0.21795	2,00E-05
7	BMI	-0.17141	0.00016
7	Hematocrit_1	-0.20108	8,00E-05
10	Gender	-0.19515	6,00E-05
10	Hematocrit_1	0.17642	3,00E-05
10	DSM_Adh_Pct	0.18715	3,00E-05
10	ASR_Rule_Raw	0.18359	1,00E-05
10	ASR_Rule_Pct	0.16697	2,00E-05
11	ProcSpeed_Unadj	-0.18789	0.0001
11	ProcSpeed_AgeAdj	-0.19773	2,00E-05
12	SSAGA_Alc_12_Frq	0.18331	0.00018
17	SSAGA_Alc_Hvy_Drinks_Per_Day	-0.20727	5,00E-05
20	Age	0.21502	1,00E-06
20	SSAGA_Rlshp	0.1886	0.00022
21	Gender	-0.20289	4,00E-05
21	Height	0.19968	0.0001
23	FearSomat_Unadj	-0.11051	0.00016
24	Weight	-0.37271	1,00E-06
24	BMI	-0.3416	1,00E-06
24	SSAGA_BMICat	-0.32902	1,00E-06
24	SSAGA_BMICatHeaviest	-0.34354	1,00E-06
24	BPSystolic	-0.18895	1,00E-05
24	Opiates	-0.18117	0.00014
24	ER40SAD	0.19129	0.00015
25	Gender	0.2001	7,00E-05
25	Height	-0.27496	1,00E-06
29	Height	0.25654	2,00E-05
29	Endurance_Unadj	0.21448	1,00E-05
29	Endurance_AgeAdj	0.20108	1,00E-05
30	PSQI_Other	0.18811	6,00E-05
38	Strength_Unadj	0.18989	0.00014
38	Strength_AgeAdj	0.20355	4,00E-05
39	BMI	-0.18183	0.00011

48	Weight	0.21292	5,00E-05
48	BMI	0.20723	5,00E-05
48	SSAGA_BMICat	0.20286	9,00E-05
52	Avg_Weekend_Hard_Liquor_7days	-0.17451	5,00E-05
55	Weight	0.23069	1,00E-06
55	BMI	0.20218	2,00E-05
55	SSAGA_BMICat	0.17109	0.00016
55	SSAGA_BMICatHeaviest	0.17351	0.00011
58	Social_Task_Random_Perc_TOM	-0.15964	0.00021
63	FearAffect_Unadj	0.18134	3,00E-05
63	NEOFAC_N	0.17647	0.00018
63	PSQI_Score	0.1634	4,00E-05
64	SSAGA_Alc_D4_Ab_Sx	-0.15831	0.00015
65	Gender	-0.18439	0.00014
68	AngAggr_Unadj	-0.16719	0.00022
68	WM_Task_0bk_Place_Acc	0.15269	0.00022
87	WM_Task_2bk_Face_Acc	0.16711	0.00018
89	Gender	-0.22058	1,00E-06
89	Hematocrit_1	0.24688	1,00E-06
89	SSAGA_Alc_12_Drinks_Per_Day	0.20449	2,00E-05
89	SSAGA_Alc_12_Frq	-0.16777	4,00E-05
89	SSAGA_Alc_12_Frq_5plus	-0.19288	1,00E-06
89	SSAGA_Alc_Hvy_Frq	-0.18465	2,00E-05
91	WM_Task_2bk_Body_Median_RT_Target	-0.13267	0.00015
96	Relational_Task_Acc	-0.15622	0.00017

Table S2: Significant results. First column presents the component number, second the behavioral or demographic measure it correlates with and third and fourth columns present the correlation value and the permutation p-value. Significance is defined at $p < 0.05$ and we used FDR correction for multiple correction ($q < 2.2 \times 10^{-4}$).

Spatial maps

In Figure S1 we report spatial maps associated with the components showing at least one Bonferroni corrected significant relationship with any behavioral measure ($q < 1.4 \times 10^{-6}$), which are not fully reported in the main text. The left column shows the independent component number. For components 1, 2 and 89 we show spatial maps for all modalities contributing to that component (Figure S2) i.e. VBM and PA for component number 1 and just VBM for components number 2 and 89. To provide a clearer interpretation we decided to show a selection of the

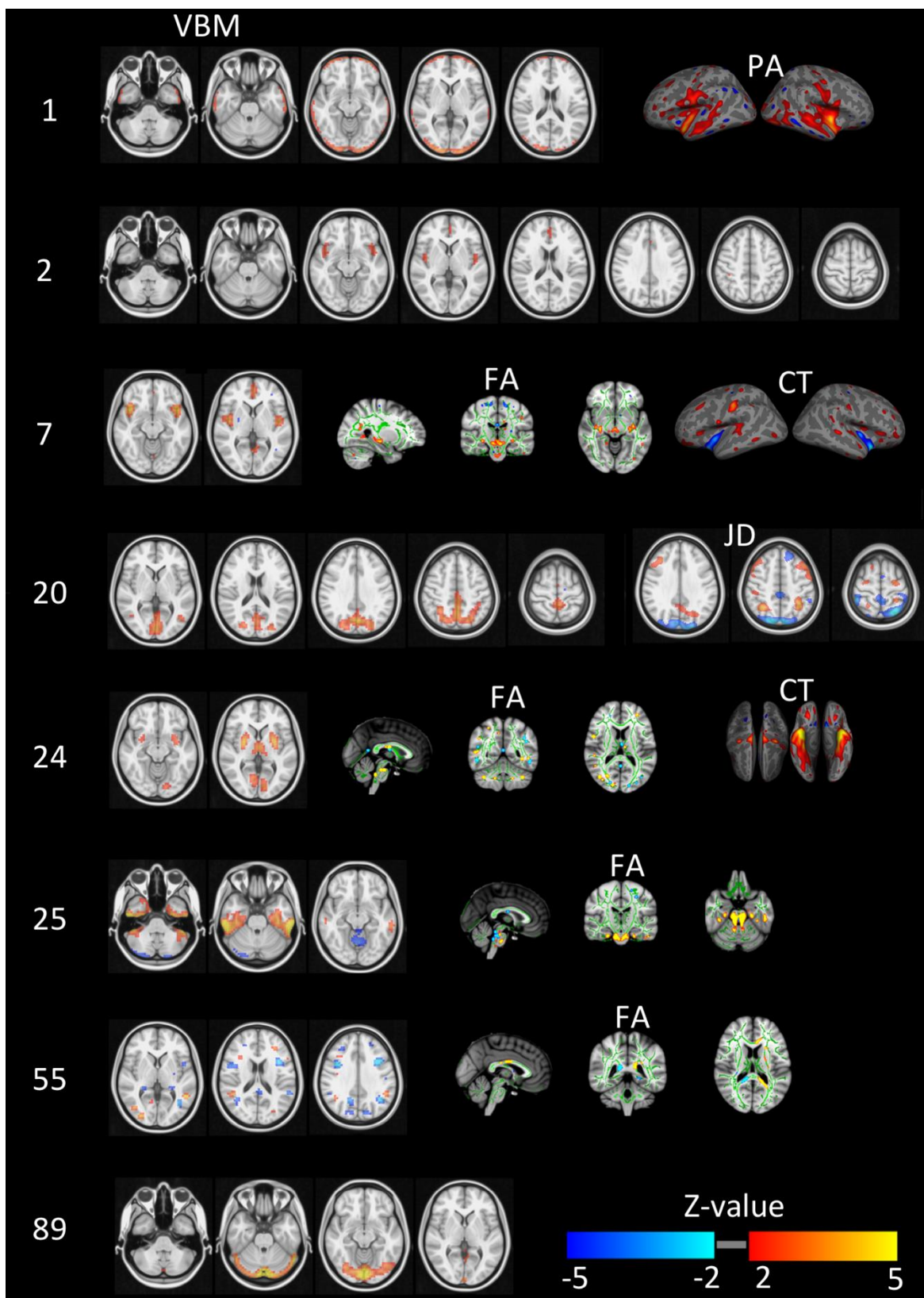


Figure S1: Summary of spatial maps associated with the components indicated in the most left column.

[Type here]

relevant modalities for the other components and full nifty maps will be separately uploaded as SI. Here, we provide VBM, FA and CT for components number 7 and number 24; VBM and JD for component number 20 and finally VBM and FA for 25, 29 and 55.

Component number 1 mainly relates to gender, physical strength and language and it is defined by voxel based morphometry and cortical area. Its associated spatial pattern appears to reflect brain size and cortical area differences in both temporal lobes. Component 2 is driven by VBM and maps into gender, age, height, weight and strength. Its spatial extent includes the paracingulate gyrus and bilateral insular and opercular cortex. Component 7 maps into gender and shows cingulate gyrus and insular cortex. Component 20 maps grey matter density in the posterior midline into age and relationship status. Component 24 maps into weight and body mass and its mapped into putamen, intracalcarine cortex and thalamus. Components 25 and 29 relate height with the inferior temporal gyrus and the cerebellum together with strong DWI weightings in the brainstem. Component 55 relates to weight and maps into the precentral gyrus and asymmetric differences in DWI measures. Component 89 maps into hematocrit and the lingual and occipital fusiform gyrus.

Feature modalities relative contribution to components

In Figure S2 we color code the relative contribution of each feature modality to the components that show at least one Bonferroni corrected significant relationship to behavior or demographics measures ($q < 1.4 \times 10^{-6}$).

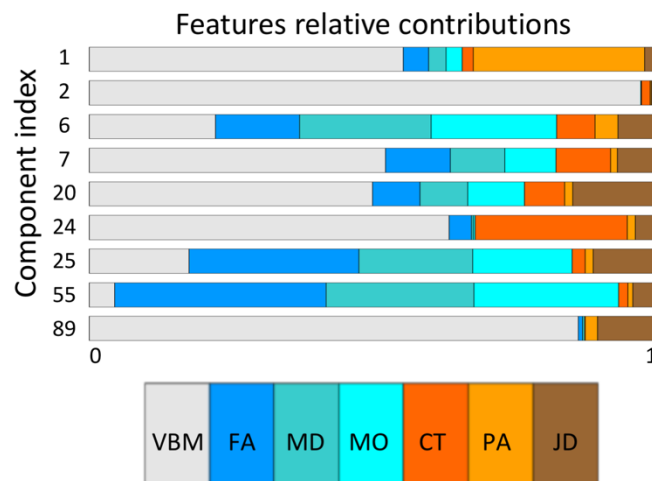


Figure S2: Relative contributions of each feature modality to the most relevant components.

Robustness: model order

In this section we assess the robustness of the results with respect to the model order choice. To that end we perform a correlation analyses between the reported 100 dimensional factorization subjects-mode and that of a 90 and a 110 dimensional factorizations. In the top row of Figure S3 we present correlation matrices between a 100 dimensional factorization (y-axis) and a 90 and 110 dimensional factorizations (top left and right panels respectively). Only significant correlations after Bonferroni correction are reported (i.e. p-value smaller than $0.05 / (100 \times 90)$ and $0.05 / (100 \times 110)$ respectively). For each component of the 100 dimensional factorization we present in the bottom row of Figure S3 the absolute correlation value with each of the components of the different dimensionality factorization. We appreciate that most components are recovered with high accuracy (r close to 1) independently of the order of the factorization. The black dashed lines represent the most relevant of the reported components, independent component number 6.

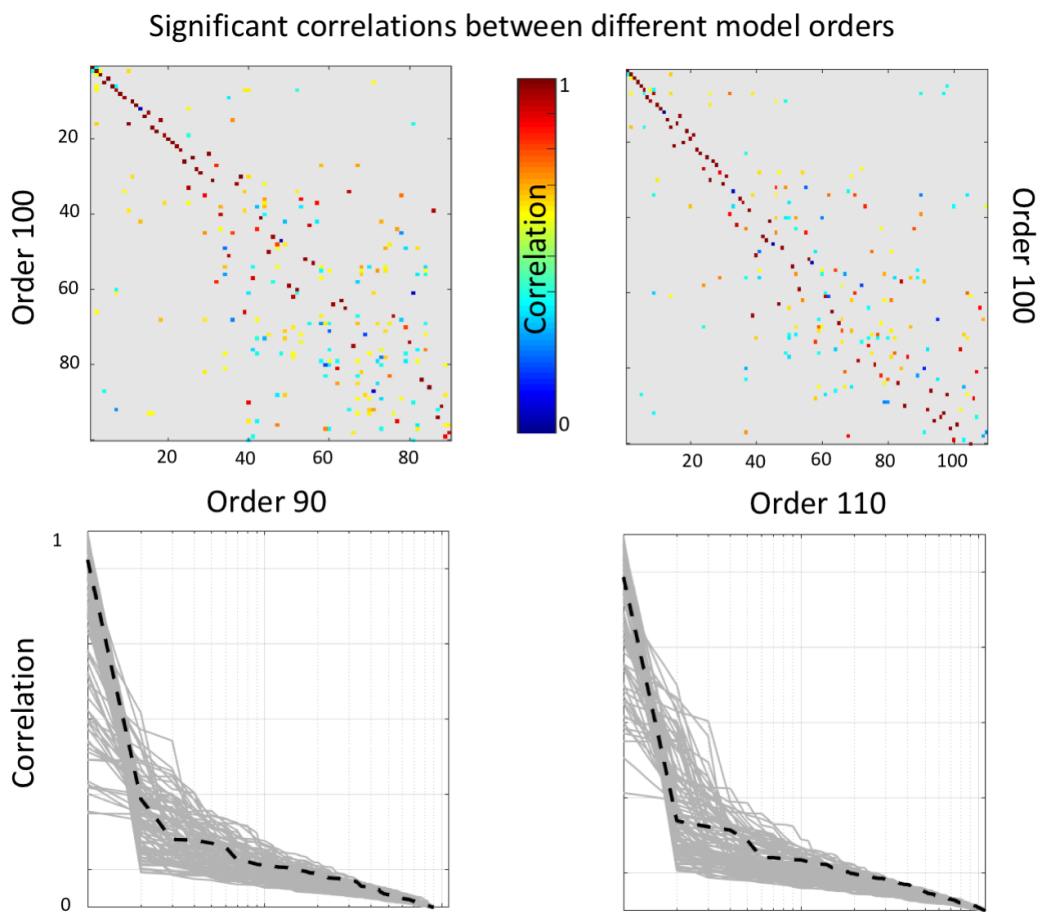


Figure S3: Top: Significant correlations between the reported (100 dimensional) factorization and a 90 dimensional (left panel) and 110 dimensional (right panel). Bottom: sorted absolute correlations for each of the components of the reported factorization with the other model orders components. The black discontinuous line represents component number 6.

Robustness: analyses without the Jacobians

We present results summarizing the significant findings when performing an analogous analysis to the one reported in the main manuscript without the use of the JD feature. In Table S3 we present a comparison between the positive-negative mode as reported in the main text, and the set of behavioral measures significantly associated to component number 9 on the new analysis.

Behavioral and demographics	Linked ICA component 6 - correlations	Linked ICA (no JD) - correlations
Relational Task	0,35	0,29
Picture Vocabulary test	0,35	0,29
Life Satisfaction	0,32	0,3
Years of Education and Income	0,3	0,3
Working Memory	0,27	0,27
Language	0,27	0,26
Picture Sequence	0,26	0,23
Fluid intelligence (correct responses)	0,25	0,24
Emotional Support	0,22	
Social Task (Random Perc. Random)	0,21	
MMSE_Score	0,21	
Gender	0,16	
Sadness	-0,15	
Perceived Stress	-0,18	
Social Task (Random Perc. Unsure)	-0,18	
BPSystolic	-0,18	
Weight and Strength	-0,19	-0,23
Loneliness	-0,20	
DSM	-0,23	-0,25
ASR	-0,23	-0,24
Family History Drugs & Alcohol	-0,25	
Fluid intelligence (skipped responses)	-0,25	-0,25
Positive test for THC (cannabis)	-0,34	-0,31

Table S3: Comparison between the positive-negative mode presented in the main text and the multi-modal analyses excluding the JD feature (right column).

Note that component number 9 in this analysis (without the JD feature) corresponds to the component number 6 reported in the main text and it recovers the strongest behavioral associations. Further, this component presented other significant relationships not appearing significant in the main reported mode, e.g. personality related measures (NEOFAC-C).

Robustness: Analyzing morphometric differences

In this section we perform an independent component analyses (ICA) factorization²⁹ only of the Jacobian determinant matrices followed by post-hoc correlation analyses with the behavioral and demographic measures. We performed a 100-dimensional factorization and found a set of 9 components significantly correlating (Bonferroni corrected) with the component number 6 obtained from the multi-modal Linked ICA analyses. In Table S4 we present the correspondence between the positive-negative mode we found through the multimodal Linked ICA analyses and these 9 components. As before, for the multimodal analyses we only report correlations to behavior significant after FDR correction and for the JD analyses we mark with double asterisk the significant relationships after FDR correction and with a single asterisk the nominal or uncorrected significant relations ($p < 0.01$). We appreciate that although sub FDR significance threshold we observe some correspondence between the purely morphometric differences and the positive-negative mode, these relationships disappear after statistical correction for multiple comparisons.

Behavioral and demographics	Linked ICA component 6 - correlations	ICA in JD (9 components) - correlations
Relational Task	0,35	0,17*
Picture Vocabulary test	0,35	0,23
Life Satisfaction	0,32	0,15*
Years of Education and Income	0,3	0,17
Working Memory	0,27	0,2
Language	0,27	0,17*
Picture Sequence	0,26	
Fluid intelligence (correct responses)	0,25	0,2*
Emotional Support	0,22	0,13*
Social Task (Random Perc, Random)	0,21	0,14*
MMSE_Score	0,21	0,12*
Gender	0,16	0,14*
Sadness	-0,15	-0,14*
Perceived Stress	-0,18	-0,15*

Social Task (Random Perc. Unsure)	-0,18	-0,11*
BPSystolic	-0,18	-0,12*
Weight and Strength	-0,19	-0,15*
Loneliness	-0,20	
DSM	-0,23	-0,09*
ASR	-0,23	-0,17*
Family History Drugs & Alcohol	-0,25	-0,17*
Fluid intelligence (skipped responses)	-0,25	-0,22*
Positive test for THC (cannabis)	-0,34	-0,17*

Table S4: Summary of the uni-modal analyses using the JD feature. In the second row all relationships are significant after multiple comparison correction. For the uni-modal analysis (the third row), significant associations after multiple comparison correction are denoted with a double asterisk and nominal significant but not significant after multiple comparison correction are marked with a single asterisk.