RUNNING TITLE: Personality and brain morphometry.

No Evidence for Associations Between the Big Five Personality Traits and Variability in Brain Gray or White Matter

Reut Avinun[1,2], Salomon Israel[2], Annchen R. Knodt[1], and Ahmad R. Hariri[1]

[1]*Laboratory of NeuroGenetics, Department of Psychology & Neuroscience, Duke University, Durham, NC, USA*

[2]*Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel*

Corresponding Author: Reut Avinun, Ph.D., Laboratory of NeuroGenetics, Department of Psychology & Neuroscience, Duke University, Gray Building 2020 West Main St, Ste 0030 Durham, NC 27705. Email: reut.avinun@duke.edu

ACKNOWLEDGEMENTS AND DISCLOSURES

We thank the Duke Neurogenetics Study participants and the staff of the Laboratory of NeuroGenetics. The Duke Neurogenetics Study received support from Duke University as well as US-National Institutes of Health grants R01DA033369 and R01DA031579. RA, ARK, and ARH received further support from US-National Institutes of Health grant R01AG049789. RA is also partly supported by a fellowship from the Jerusalem Brain Community. The Brain Imaging and Analysis Center was supported by the Office of the Director, National Institutes of Health under Award Number S10OD021480. The authors declare no competing financial or other interests.

Abstract

Attempts to link the Big Five personality traits of Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism with variability in trait-like features of brain structure have produced inconsistent results. Small sample sizes and heterogeneous methodology have been suspected in driving these inconsistencies. Here, we tested for associations between the Big Five personality traits and multiple measures of brain structure using data from 1,107 university students (636 women, mean age 19.69±1.24 years) representing the largest attempt to date. In addition to replication analyses based on a prior study, we conducted exploratory whole-brain analyses. Three supplementary analyses were also conducted to examine 1) possible associations with lower-order facets of personality; 2) modulatory effects of sex; and 3) effect of controlling for non-target personality traits. The analyses failed to identify any significant associations between the Big Five personality traits and variability in measures of cortical thickness, surface area, subcortical volume, or white matter microstructural integrity. The supplementary analyses supported these primary null findings, suggesting they were not substantively biased by our choice of analytic model. These results collectively indicate that if there are direct associations between the Big Five personality traits and variability in brain structure they are of likely very small effect sizes requiring very large samples for reliable detection.

Introduction

Studies regarding the basic structure of individual differences in personality traits have yielded a relatively consistent five factor model (FFM), which includes Openness-to-Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (Digman, 1990). In contrast, studies of how these so-called Big Five personality traits relate to underlying trait-like features of the brain have yet to identify consistent patterns despite a growing number of attempts (e.g., Bjørnebekk et al., 2013; Coutinho et al., 2013; DeYoung et al., 2010; Ferschmann et al., 2018; Hu et al., 2011; Kapogiannis et al., 2013; Liu et al., 2013; Lu et al., 2014). Unlike the studies from which the Big Five were derived, most neuroimaging studies of personality traits have relied on relatively small samples (N<100). In fact, a recent study has suggested that even samples of 300 participants may be too small to reliably detect associations between psychological phenotypes and brain morphometry (Kharabian Masouleh et al., 2019). Consequently, consistent and replicable links have yet to emerge (reviewed in Allen and DeYoung, 2017; Yarkoni, 2015). This lack of statistical power has been further compounded by varied methodological and analytic approaches across studies.

Here, we tested for associations between the Big Five personality traits and multiple features of brain structure in the largest sample to date (N=1,107). Notably, other than its size, our sample also had the advantage of being relatively homogeneous in age (18-22 years), which may affect associations between personality traits and brain structure (Ferschmann et al., 2018). The specific measures of brain gray matter were cortical thickness (CT), surface area (SA), and

subcortical volume. Fractional anisotropy (FA) was used as a specific measure of white matter microstructural integrity.

We examined associations with cortical thickness (CT) and surface area (SA) separately rather than the courser measure of gray matter volume, which is the product of these two measures. Specific associations between the Big Five personality traits and either CT or SA would be informative because these structural measures exhibit different developmental trajectories (Wierenga et al., 2014) and distinct genetic influences (Panizzon et al., 2009). Relatedly, we conducted surface-based parcellation analyses rather than traditional vertex- or voxel-based analyses to maximize spatial resolution otherwise lost to smoothing across tissue types (i.e., CSF, gray matter, and white matter) and anatomical regions (Coalson et al., 2018; Glasser et al., 2016). A parcellation-based approach further restricts the number of tests conducted to anatomically defined regions, thereby minimizing the number of comparisons conducted. We also explicitly controlled for race/ethnicity, because previous research has found links between these variables and brain structure (e.g., Brickman et al., 2008; Pfefferbaum et al., 2016; Xie et al., 2015), as well as personality (Foldes et al., 2008). Lastly, we used regression analyses with robust standard errors that accommodate non-normality in data, because previous research has indicated that the CT and SA of some brain regions may not be normally distributed (Patel et al., 2018).

Using the above general strategy, we conducted three related sets of analyses. First, we attempted to replicate the personality associations with CT and SA reported by Hyatt et al. (2019), which represents the largest previously published study of the structural brain correlates of personality based on data from 1,104 twins and siblings

from the Human Connectome Project (age range: 22-36). Second, we conducted whole-brain exploratory analyses to examine all possible associations between the Big Five personality traits and SA, CT, subcortical volume, and FA. Third, because previous studies suggest a) that lower order facets of the Big Five personality traits may better correspond with brain structure (Bjørnebekk et al., 2013); b) associations may differ by sex (Nostro et al., 2016); and c) associations may only be present when controlling for non-target traits (e.g., DeYoung et al., 2010; Liu et al., 2013; Riccelli et al., 2017), we conducted three supplementary analyses to address these possibilities and further inform future research in personality neuroscience.

Methods

Participants

1330 participants (762 women, mean age 19.70±1.25 years) successfully completed the Duke Neurogenetics Study (DNS), which assessed a range of behavioral and biological traits among young adult, university students. The DNS was approved by the Duke University School of Medicine Institutional Review Board, and all participants provided written informed consent prior to participation. All participants were free of the following study exclusions: 1) medical diagnoses of cancer, stroke, diabetes requiring insulin treatment, chronic kidney or liver disease, or lifetime history of psychotic symptoms; 2) use of psychotropic, glucocorticoid, or hypolipidemic medication; and 3) conditions affecting cerebral blood flow and metabolism (e.g., hypertension). Importantly, neither current nor lifetime diagnosis were an exclusion criterion, as the DNS sought to establish broad variability in multiple behavioral phenotypes related to psychopathology.

The current analyses of gray matter (i.e., CT, SA, and subcortical volume) were conducted on a subset of 1107 participants (636 women, mean age 19.69±1.24 years) for whom there was T1-weighted structural imaging data available post quality control procedures (see below) as well as personality questionnaire and genetic race/ethnicity data. Amongst this subset, 224 participants had at least one DSM-IV diagnosis. Based on self-report, there were 499 non-Hispanic Caucasians, 125 African Americans, 294 Asians, 71 Latino/as, 2 Pacific Islanders, and 116 multiracial or other participants in this subset.

White matter microstructure analyses were conducted on a further subset of 778 participants (443 women, mean age 19.67±1.25 years) for whom there was diffusion weighted imaging data available post quality control procedures (see below) as well as personality questionnaire and genetic race/ethnicity data. Amongst this subset, 156 participants had at least one DSM-IV diagnosis. Based on self-report, there were 351 non-Hispanic Caucasians, 92 African Americans, 213 Asians, 47 Latino/as, 2 Pacific Islanders, and 73 multiracial or other participants in this subset.

Race/Ethnicity

Because self-reported race and ethnicity are not always an accurate reflection of genetic ancestry, an analysis of identity by state of whole-genome SNPs was performed in PLINK (Purcell et al., 2007). The first four multidimensional scaling components were used as covariates to reduce possible confounding effects of race/ethnicity. The decision to use only the first four components was based on an

examination of a scree plot of eigenvalue, which showed that the eigenvalues became very similar after the fourth component.

Personality

The NEO personality inventory revised (NEO-PI-R; Costa and McCrae, 2008), which includes 240 items, was used to assess the FFM of personality. Each personality trait consists of 6 lower order facets, which were modeled in our supplementary analyses. The Cronbach's alpha of the Big Five ranged between .70 to .85.

MRI Data Acquisition

Each participant was scanned using one of two identical research-dedicated GE MR750 3T scanners at the Duke-UNC Brain Imaging and Analysis Center. Each identical scanner was equipped with high-power high-duty cycle 50-mT/m gradients at 200 T/m/s slew rate and an eight-channel head coil for parallel imaging at high bandwidth up to 1 MHz. T1-weighted images were obtained using a 3D Ax FSPGR BRAVO sequence with the following parameters: TR = 8.148 ms; TE = 3.22 ms; 162 axial slices; flip angle, 12°; FOV, 240 mm; matrix =256×256; slice thickness = 1 mm with no gap (voxel size 0.9375×0.9375×1 mm); and total scan time = 4 min and 13 s. Following an ASSET calibration scan, two 2-min 50-s diffusion imaging acquisitions were collected, providing full brain coverage with 2-mm isotropic resolution and 15 diffusion weighted directions (10-s repetition time, 84.9 ms echo time, b value 1,000 s/mm2, 240 mm field of view, 90° flip angle, 128×128 acquisition matrix, slice thickness=2 mm).

MRI Data Processing

To generate regional measures of brain morphometry, anatomical images for each subject were first skull-stripped using ANTs (Klein et al., 2009), then submitted to Freesurfer's (Version 5.3) recon-all with the "-noskullstrip" option (Dale et al., 1999; Fischl et al., 1999), using an x86_64 linux cluster running Scientific Linux. Of the 1321 participants who completed the high-resolution T1-weighted imaging protocol, 11 were excluded for the presence of motion-related or external artifacts, 4 were excluded for incidental findings, and 1 was unable to be processed with FreeSurfer. Additionally, the gray and white matter boundaries determined by recon-all were visually inspected using FreeSurfer QA Tools (https://surfer.nmr.mgh.harvard.edu/fswiki/QATools). This revealed small to moderate errors in gray matter boundary detection in 51 individuals who were consequently excluded. CT and SA for 31 regions in each hemisphere, as defined by the Desikan-Killiany-Tourville atlas (Klein and Tourville, 2012), a modified version of the Desikan-Killiany atlas (Desikan et al., 2006) meant to make region definitions as unambiguous as possible and define boundaries best suited to FreeSurfer's classifier algorithm, were extracted using Freesurfer. Additionally, gray matter volumes from eight subcortical regions (Cerebellum Cortex, Thalamus, Caudate, Putamen, Pallidum, Hippocampus, Amygdala, and Accumbens area) were extracted with Freesurfer's subcortical segmentation ("aseg") pipeline (Fischl et al., 2002). Estimated Total Intracranial Volume (ICV), total gray matter volume, cerebral white matter volume, and left and right hemisphere mean CT were also extracted from the "aseg" pipeline, and average whole-brain CT was calculated based on the estimates for the left and right hemispheres.

Diffusion weighted images were processed according to the Diffusion Tensor Imaging (DTI) protocol developed by the Enhancing Neuro Imaging Genetics Through Meta-Analysis (ENIGMA) consortium (Jahanshad et al., 2013; or http://enigma.ini.usc.edu/protocols/dti-protocols/). In brief, raw diffusion-weighted images underwent eddy current correction and linear registration to the non-diffusion weighted image in order to correct for head motion. These images were skull-stripped and diffusion tensor models were fit at each voxel using FMRIB's Diffusion Toolbox in FSL (FDT; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT), and the resulting two fractional anisotropy (FA) maps were linearly registered to each other and then averaged. Average FA images from all subjects were non-linearly registered to the ENIGMA-DTI target FA map, a minimal deformation target calculated across a large number of individuals (Jahanshad et al., 2013). The images were then processed using the tract-based spatial statistics (TBSS) analytic method (Smith et al., 2006) modified to project individual FA values onto the ENIGMA-DTI skeleton. Following the extraction of the skeletonized white matter and projection of individual FA values, tract-wise regions of interest, derived from the Johns Hopkins University (JHU) white matter parcellation atlas (Mori et al., 2005), were transferred to extract the mean FA across the full skeleton and average whole-brain FA values for a total of 24 (partially overlapping) regions across the two scans. All FA measures from the right and left hemispheres were averaged. Additionally, volume-by-volume head motion was quantified by calculating the root mean square (RMS) displacement of the six motion parameters (three translation and three rotation components), for each pair of consecutive diffusion-weighted brain volumes. The resulting volume-by-volume RMS deviation values were averaged across all images,

yielding a summary statistic of head motion for each participant to add to the FA analyses as a covariate, as previously recommended for DTI analyses (Yendiki et al., 2014).

Statistical Analyses

Analyses were conducted in R version 3.5.1 (R Core Team, 2018), with the packages "broom" (Robinson and Hayes, 2018), "tidyr" (Wickham and Henry, 2018), "dplyr" (Wickham et al., 2019), "lmtest" (Zeileis and Hothorn, 2002), "readr" (Wickham et al., 2018), and "sandwich" (Zeileis, 2004). Linear regression analyses with robust standard errors were performed with brain measures as outcomes, personality measures as independent variables, and sex, age, and four ancestry-informative genetic principal components as covariates of no interest. ICV, average CT, and average FA, were used as additional covariates for analyses of subcortical volume and surface area, CT, and FA, respectively. For the FA analyses, which can be particularly sensitive to motion, head motion was also included as a covariate. The Big Five personality traits were standardized (M=0, SD=1) in SPSS version 25 before analyses.

We first attempted to replicate the personality and SA/CT significant associations reported by Hyatt et al. (2019; Table 1) at $p<.005$ (i.e., the findings that were reported as significant in their paper). We next proceeded to conduct exploratory parcellation-based analyses across the whole-brain (31 SA regions, 31 CT regions, 8 subcortical regions, 24 FA measures, and total gray matter, cerebral white matter, total SA, and mean CT) for each of the Big Five personality traits (a total of

5*98=490 tests). The "false discovery rate" (FDR) adjustment (Benjamini and Hochberg, 1995) was applied to correct for multiple comparisons with the p.adjust function in R. Lastly, we conducted three supplementary analyses: 1) whole-brain exploratory parcellation-based analyses of the Big Five personality facets; 2) whole-brain exploratory parcellation-based analyses of the Big Five personality traits for men and women, separately; and 3) whole-brain exploratory parcellation-based analyses of each Big Five personality trait while controlling for the other four traits.

Results

Replication of Hyatt et al. (2019)

As reported in Table 2, none of the 14 associations that were significant at p<.005 in Hyatt et al. (2019) were significant in our analyses, even without correcting for multiple comparisons (i.e., using an uncorrected p<.05 threshold). As Hyatt et al. did not control for race/ethnicity, we also ran analyses without the genetic principal components to test whether these could account for the different results. Again, none of the associations remained significant after correcting for multiple comparisons, but two associations were significant at an uncorrected p<.05: a positive association between neuroticism and CT of the superior frontal gyrus (b=0.0069, SE=0.0027, p=.009) and a positive association between neuroticism and the SA of the right supramarginal gyrus (b=23.803, SE=11.455, p=.03). As the race/ethnic composition of our sample differed from the race/ethnic composition of the Human Connectome Project (HCP) sample, on which the analyses of Hyatt et al. were based (i.e., 44.6% vs. 74.8% of non-Hispanic Caucasians, 11.4% vs. 15.1% of African-Americans, and 27% vs. 5.7% of Asian-Americans, which also included Native

Hawaiian, or other Pacific Islander in the HCP), we also separately present the results from our three largest ethnic subsamples, as determined based on self-reports and genetic ancestry components, when available. Here, the sample sizes are larger because individuals with missing genetic data were also included based on self-reported race/ethnicity: non-Hispanic Caucasians (n=559), Asians (n=336), and African-Americans (n=143). As reported in Table 2, there were differences in the direction of the effect between the groups, supporting our decision to control for race/ethnicity (e.g., the association between extraversion and the SA of the superior frontal gyrus was negative in Asian-Americans and positive in non-Hispanic Caucasians and African-Americans).

Exploratory whole-brain analyses

There were no significant associations between the Big Five personality traits and CT, SA, subcortical volume, or FA after FDR correction for multiple comparisons. The top associations (i.e., uncorrected p<.005) are reported in Table 3. All associations are further presented in Supplementary Table 1.

Supplementary analyses

Our supplementary analyses (i.e., testing personality facets instead of the Big Five personality traits, conducting sex-specific analyses, or using non-target personality traits as covariates; reported in Supplementary Tables 2-5) revealed that these null findings were not biased by our choice of analytic model. The only association that survived FDR correction for multiple comparisons across all the tests conducted in

the current study (N=4750) was between the SA of the superior temporal gyrus and the dutifulness facet of conscientiousness (b=-.39.65, SE=8.77, p=6.84e-06, FDR adjusted p=0.033), such that higher dutifulness predicted reduced superior temporal gyrus SA. The R^2 of the regression model without dutifulness was 0.6537, and the R^2 of the model with dutifulness was 0.6596, indicating that dutifulness was associated with 0.6% of the variance in the SA of the superior temporal gyrus.

Discussion

In the current study, with the largest sample to date, we failed to identify robust links between the Big Five personality traits and multiple, trait-like features of brain structure. Several supplementary analyses, in which we tested the facets of personality, conducted separate analyses for men and women, and included the non-target personality traits as covariates, confirmed these primary null findings. There was one exception: an association between greater SA of the superior temporal gyrus, an area involved in language perception and production, and lower scores on the dutifulness facet of conscientiousness. However, Bjørnebekk et al. (2013) who reported an association with the caudal part of the superior temporal gyrus, did so for scores on different facets of conscientiousness (achievement striving and self-discipline), and this association was not replicated in other studies (Hyatt et al., 2019; Nostro et al., 2016). Consequently, this singular association in our current analyses should be treated with caution. Generally, the supplementary analyses suggested that the null results of the primary analyses did not depend on specific analytical or methodological choices.

There are several possible reasons for the lack of replicable associations in mapping personality and brain structure including the current failure to replicate associations identified by Hyatt et al. from analyses of the previously largest sample. With regard to this specific failure, it is possible that unaccounted-for effects of family structure in the data derived primarily from twins and siblings (Van Essen et al., 2013), biased their observed associations. More generally, most of the correlations reported by Hyatt et al. were smaller than .1, and in our study the only significant association explained 0.6% of the variance. These results suggest that the effect sizes of direct associations between personality traits and brain structure are likely to be very small and will require very large sample sizes to be reliably detected.

Furthermore, differences between and within samples may also limit replicability in personality neuroscience. Age is known to affect brain structure, and indeed has been shown to moderate associations between brain structure and personality (Ferschmann et al., 2018). Sex differences may also be relevant as has been shown by Nostro et al., (2016) and in our supplementary results, where, although an interaction by sex was not formally tested, the top associations differed between men and women. Our results also suggest that accounting for race/ethnicity may be required as well when testing for personality-brain structure associations. Indeed, previous research has shown differences in brain structure as a function of race/ethnicity (e.g., Brickman et al., 2008; Pfefferbaum et al., 2016; Xie et al., 2015). For example, a different brain atlas than the one constructed based on Caucasians may be needed to accurately identify variability in structure from a different racial/ethnic population (Tang et al., 2010). Thus, it may be insufficient to simply control for race/ethnicity in analytic models.

Although we used the largest sample to date, included 260 items to assess personality, and employed different methodologies to test for the associations between the Big Five personality traits and brain structure, our study does have several limitations. First, we did not exhaust all the possible ways to assess personality. For example, an alternative classification approach is represented by personality "types," which defines categories of individuals based on similar configurations of interacting traits. A large analysis of personality types indicated that there are 4 personality types that can be clustered based on scores on the Big Five personality traits (Gerlach et al., 2018). Thus, for example, someone low on neuroticism may also have an average or a high conscientiousness score, which may correspond differently with brain structure. Future studies could focus on such "types" in examining personality-brain structure associations. Second, our acquisition protocol precluded the application of more anatomically precise parcellation schemes (e.g., Glasser et al., 2016). Third, our sample of volunteer students at a top university may not be representative of the general population. Lastly, we did not examine brain function. The brain correlates of personality may be more readily identified in functional measures, such as functional connectivity. However, as functional MRI studies suffer from small sample sizes, here as well caution will be needed in the interpretation of findings until replicable findings emerge in large samples.

Our null findings echo comments made by Yarkoni (2015): "There is no guarantee that any particular psychometric model of individual differences in personality will map onto underlying biological process models in any straightforward way. In fact…a clear-cut relationship between the two is likely to be

the exception rather than the rule." Additionally, a recent replicability study concluded that associations between psychological measures (including personality) and specific brain structures in a healthy sample are "highly unlikely" (Kharabian Masouleh et al., 2019). Our results are consistent with this bleak perspective. That said, small effect sizes and possible moderating effects of sex, age, and race/ethnicity suggest the possibility that with ever larger and more homogeneous samples reliable links between personality and trait-like features of the brain may yet emerge. Here, the field of personality neuroscience may benefit from following the lead of genome-wide association studies that have, after many failed attempts with candidate gene studies and small samples (e.g., Avinun et al., 2018; Bosker et al., 2011), begun to reveal the genetic architecture of complex traits through massive samples (Plomin and von Stumm, 2018). The growth of shared imaging data through research consortia (e.g., the "enhancing neuroimaging genetics through meta-analysis" project [ENIGMA]) may allow for such gains in personality neuroscience sooner than later.

References

Allen, T.A., DeYoung, C.G., 2017. Personality neuroscience and the five factor model. In: Widiger, T.A. (Ed.), Oxford handbook of the five factor model. Oxford University Press, New York, pp. 319-352.

Avinun, R., Nevo, A., Knodt, A.R., Elliott, M.L., Hariri, A.R., 2018. Replication in Imaging Genetics: The Case of Threat-Related Amygdala Reactivity. Biological Psychiatry 84, 148-159.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) 57, 289-300.

Bjørnebekk, A., Fjell, A.M., Walhovd, K.B., Grydeland, H., Torgersen, S., Westlye, L.T., 2013. Neuronal correlates of the five factor model (FFM) of human personality: Multimodal imaging in a large healthy sample. Neuroimage 65, 194-208.

Bosker, F., Hartman, C., Nolte, I., Prins, B., Terpstra, P., Posthuma, D., Van Veen, T., Willemsen, G., DeRijk, R., De Geus, E., 2011. Poor replication of candidate genes for major depressive disorder using genome-wide association data. Molecular Psychiatry 16, 516.

Brickman, A.M., Schupf, N., Manly, J.J., Luchsinger, J.A., Andrews, H., Tang, M.X., Reitz, C., Small, S.A., Mayeux, R., DeCarli, C., 2008. Brain morphology in older African Americans, Caribbean Hispanics, and whites from northern Manhattan. Archives of Neurology 65, 1053-1061.

Coalson, T.S., Van Essen, D.C., Glasser, M.F., 2018. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. Proceedings of the National Academy of Sciences 115, E6356-E6365.

Costa, P.T., McCrae, R.R., 2008. The revised neo personality inventory (neo-pi-r). SAGE, Thousands Oaks, CA.

Coutinho, J.F., Sampaio, A., Ferreira, M., Soares, J.M., Gonçalves, O.F., 2013. Brain correlates of pro-social personality traits: a voxel-based morphometry study. Brain Imaging and Behavior 7, 293-299.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9, 179-194.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968-980.

DeYoung, C.G., Hirsh, J.B., Shane, M.S., Papademetris, X., Rajeevan, N., Gray, J.R., 2010. Testing predictions from personality neuroscience: Brain structure and the big five. Psychological Science 21, 820-828.

Digman, J.M., 1990. Personality structure: Emergence of the five-factor model. Annual Review of Psychology 41, 417-440.

Ferschmann, L., Fjell, A.M., Vollrath, M.E., Grydeland, H., Walhovd, K.B., Tamnes, C.K., 2018. Personality traits are associated with cortical development across adolescence: a longitudinal structural MRI study. Child Development 89, 811-822.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341-355.

Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. Neuroimage 9, 195-207.

Foldes, H.J., Duehr, E.E., Ones, D.S., 2008. Group differences in personality: Meta-analyses comparing five US racial groups. Personnel Psychology 61, 579-616.

Gerlach, M., Farb, B., Revelle, W., Amaral, L.A.N., 2018. A robust data-driven approach identifies four personality types across four large data sets. Nature human behaviour 2, 735.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171.

Hu, X., Erb, M., Ackermann, H., Martin, J.A., Grodd, W., Reiterer, S.M., 2011. Voxel-based morphometry studies of personality: Issue of statistical model specification—effect of nuisance covariates. Neuroimage 54, 1994-2005.

Hyatt, C.S., Owens, M.M., Gray, J.C., Carter, N.T., MacKillop, J., Sweet, L.H., Miller, J.D., 2019. Personality traits share overlapping neuroanatomical correlates with internalizing and externalizing psychopathology. Journal of Abnormal Psychology 128, 1.

Jahanshad, N., Kochunov, P.V., Sprooten, E., Mandl, R.C., Nichols, T.E., Almasy, L., Blangero, J., Brouwer, R.M., Curran, J.E., de Zubicaray, G.I., 2013. Multi-site

genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA–DTI working group. Neuroimage 81, 455-469.

Kapogiannis, D., Sutin, A., Davatzikos, C., Costa Jr, P., Resnick, S., 2013. The five factors of personality and regional cortical variability in the Baltimore longitudinal study of aging. Human Brain Mapping 34, 2829-2840.

Kharabian Masouleh, S., Eickhoff, S., Hoffstaedter, F., Genon, S., 2019. Empirical examination of the replicability of associations between brain structure and psychological variables. eLife.

Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage 46, 786-802.

Klein, A., Tourville, J., 2012. 101 labeled brain images and a consistent human cortical labeling protocol. Frontiers in Neuroscience 6, 171.

Liu, W.-Y., Weber, B., Reuter, M., Markett, S., Chu, W.-C., Montag, C., 2013. The Big Five of Personality and structural imaging revisited: a VBM–DARTEL study. Neuroreport 24, 375-380.

Lu, F., Huo, Y., Li, M., Chen, H., Liu, F., Wang, Y., Long, Z., Duan, X., Zhang, J., Zeng, L., 2014. Relationship between personality and gray matter volume in healthy young adults: a voxel-based morphometric study. PLoS ONE 9, e88763.

Mori, S., Wakana, S., Van Zijl, P.C., Nagae-Poetscher, L., 2005. MRI atlas of human white matter. Elsevier, Amsterdam.

Nostro, A.D., Müller, V.I., Reid, A.T., Eickhoff, S.B., 2016. Correlations between personality and brain structure: a crucial role of gender. Cerebral Cortex 27, 3698-3712.

Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Jernigan, T.L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M.J., Grant, M.D., Franz, C.E., 2009. Distinct genetic influences on cortical surface area and cortical thickness. Cerebral Cortex 19, 2728-2735.

Patel, S., Patel, R., Park, M.T.M., Masellis, M., Knight, J., Chakravarty, M.M., 2018. Heritability estimates of cortical anatomy: the influence and reliability of different estimation strategies. Neuroimage 178, 78-91.

Pfefferbaum, A., Rohlfing, T., Pohl, K.M., Lane, B., Chu, W., Kwon, D., Nolan Nichols, B., Brown, S.A., Tapert, S.F., Cummins, K., 2016. Adolescent development of cortical and white matter structure in the NCANDA sample: role of sex, ethnicity, puberty, and alcohol drinking. Cerebral Cortex 26, 4101-4121.

Plomin, R., von Stumm, S., 2018. The new genetics of intelligence. Nature Reviews Genetics.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics 81, 559-575.

R Core Team, 2018. R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria.

Riccelli, R., Toschi, N., Nigro, S., Terracciano, A., Passamonti, L., 2017. Surface-based morphometry reveals the neuroanatomical basis of the five-factor model of personality. Social Cognitive and Affective Neuroscience 12, 671-684.

Robinson, D., Hayes, A., 2018. broom: Convert Statistical Analysis Objects into Tidy Tibbles. R package version 0.5.1.

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487-1505.

Tang, Y., Hojatkashani, C., Dinov, I.D., Sun, B., Fan, L., Lin, X., Qi, H., Hua, X., Liu, S., Toga, A.W., 2010. The construction of a Chinese MRI brain atlas: a morphometric comparison study between Chinese and Caucasian cohorts. Neuroimage 51, 33-41.

Wickham, H., François, R., Henry, L., Müller, K., 2019. dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1.

Wickham, H., Henry, L., 2018. tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions. R package version 0.8.2.

Wickham, H., Hester, J., Francois, R.F., 2018. readr: Read Rectangular Text Data. R package version 1.3.1.

Wierenga, L.M., Langen, M., Oranje, B., Durston, S., 2014. Unique developmental trajectories of cortical thickness and surface area. Neuroimage 87, 120-126.

Xie, W., Richards, J.E., Lei, D., Lee, K., Gong, Q., 2015. Comparison of the brain development trajectory between Chinese and US children and adolescents. Frontiers in systems neuroscience 8, 249.

Yarkoni, T., 2015. Neurobiological substrates of personality: A critical overview. In: Mikulincer, M.S., Cooper, P.R., Lynne, L.M., Randy, J. (Eds.), APA handbooks in psychology. APA handbook of personality and social psychology. American Psychological Association, Washington D.C., pp. 61-83.

Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., Fischl, B., 2014. Spurious group differences due to head motion in a diffusion MRI study. Neuroimage 88, 79-90.

Zeileis, A., 2004. Econometric computing with HC and HAC covariance matrix estimators.

Zeileis, A., Hothorn, T., 2002. Diagnostic checking in regression relationships. R news 2, 7-10.

Table 1. Significant personality and brain structure (SA and CT) associations reported by Hyatt et al. 2019.

	CT	**SA**
Neuroticism	Left caudal middle frontal gyrus (+) Pars orbitalis (+) Pars triangularis (+) Superior frontal gyrus (+)	Left cuneus (-) pars triangularis (-) superior parietal lobule (-) Right supramarginal gyrus (-) Left and right superior frontal gyrus (-)
Openness	Left rostral middle frontal gyrus (-) superior parietal lobule (-)	Left inferior temporal gyrus (+)
Agreeableness	Left caudal middle frontal gyrus (-)	
Extraversion		Right superior frontal gyrus (+)

Note. +/- indicate the direction of the associations (i.e., positive or negative respectively)

Table 2. Standardized parameter estimates and standard errors from our attempt to replicate the associations reported by Hyatt et al. (2019) in all participants and within the three largest racial/ethnic subgroups from the Duke Neurogenetics Study.

	All (N=1107) b (SE)	**Caucasians (N=559)** b (SE)	**African-Americans (N=143)** b (SE)	**Asians-Americans (N=336)** b (SE)
Left cuneus (SA) on Neuroticism	-5.04391 (7.6354)	5.98811 (11.3887)	-25.57213 (18.0369)	-9.15169 (12.4039)
Pars triangularis (SA) on Neuroticism	4.08199 (5.6707)	5.17454 (8.5867)	19.93103 (16.4172)	-1.44526 (9.5138)
Superior parietal lobule (SA) on Neuroticism	11.91824 (11.2085)	5.04335 (15.8018)	37.99563 (31.5336)	-1.56423 (20.3665)
Right supramarginal gyrus (SA) on Neuroticism	19.56767 (12.3316)	12.8792 (18.3712)	13.41091 (39.9996)	22.12617 (19.7336)
Superior frontal gyrus (SA) on Neuroticism	5.89661 (16.0535)	13.2761 (22.3406)	23.02461 (46.9883)	9.75734 (25.4849)
Left inferior temporal gyrus (SA) on Openness	14.18753 (11.0602)	28.98508 (15.2697)	57.54226 (29.6909)	-9.53206 (20.0009)
Right superior frontal gyrus (SA) on Extraversion	19.56087 (18.2997)	4.59021 (26.8022)	73.67606 (55.5483)	-58.28776 (34.5373)
Left caudal middle frontal gyrus (CT) on Neuroticism	0.00336 (0.0045)	0.00629 (0.0057)	0.00012 (0.0148)	0.00535 (0.0078)
Pars orbitalis (CT) on Neuroticism	-0.0065 (0.0052)	-0.01619* (0.0066)	-0.00209 (0.0201)	-0.00206 (0.009)
Pars triangularis (CT) on Neuroticism	-0.00051 (0.0034)	-0.00292 (0.0048)	-0.01458 (0.0086)	0.00517 (0.0056)
Superior frontal gyrus (CT) on Neuroticism	0.00521 (0.0029)	0.0004 (0.004)	-0.00311 (0.0087)	0.01373** (0.0045)
Left rostral middle frontal gyrus (CT) on Openness	-0.00227 (0.0036)	0.00013 (0.0051)	-0.01208 (0.0118)	0.00203 (0.0066)
Superior parietal lobule (CT) on Openness	-0.00193 (0.0024)	0.00035 (0.003)	-0.00399 (0.0062)	-0.00265 (0.0045)
Left caudal middle frontal gyrus (CT) on Agreeableness	-0.00159 (0.0043)	-0.00542 (0.0055)	0.01237 (0.0142)	-0.00155 (0.0077)

Note. CT=cortical thickness; SA=surface area; All=entire sample, controlling for 4 genetic principal components. In all analyses, sex, age and either mean CT or intracranial volume were included as covariates. *uncorrected p<.05, **uncorrected p<.01.

Table 3. Top results (p<.005) from whole-brain exploratory analyses between the Big Five personality traits and structural brain measures in the Duke Neurogenetics Study.

	b	**SE**	**p value**	**FDR adjusted p value**
Superior temporal gyrus (SA) on Conscientiousness	-33.1612	8.622786	0.000127	0.062272
Thalamus Proper on Conscientiousness	-55.3455	15.67952	0.000433	0.10611
Postcentral gyrus (CT) on Openness	0.00718	0.002239	0.001379	0.225262
Cerebral peduncle (FA) on Neuroticism	-0.00163	0.000544	0.002835	0.347291
Transverse temporal gyrus (SA) on Conscientiousness	-3.75527	1.330542	0.004853	0.475613