7-5-2023

Brain Structure Correlates of Obsessive-Compulsive Personality Traits

Allison Moreau
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Psychological and Brain Sciences

Dissertation Examination Committee:
Ryan Bogdan, Chair
Deanna Barch
Patrick Hill
Thomas Oltmanns
Douglas Samuel

Brain Structure Correlates of Obsessive-Compulsive Personality Traits

by
Allison Moreau

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2023
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

First, I would like to thank my advisor, Ryan Bogdan, for your support and mentorship the past six years. You have put in an enormous amount of work on countless drafts over the years, helping make my research the best it can be. I would also like to thank Deanna Barch and Tom Oltmanns for serving on my both my dissertation and master's committee. Your insights and suggestions throughout my graduate training have helped hone my ability to think critically about my research. Additionally, I would like to thank Doug Samuel for his OCPD expertise throughout the project and generously providing the external validation dataset that greatly strengthened this project. I also very much appreciate Patrick Hill serving on my dissertation committee and providing your thoughts and insights on the project.

I also want to acknowledge two of my undergraduate professors at Wheaton College, Grace Baron and Bianca Cody Murphy, who were formative role models in my journey to becoming a psychologist. You took me under your wing not only professionally but also personally, which I am forever grateful for. This gave me the confidence to pursue a Ph.D. in clinical psychology. I wish Dr. Baron was still with us to see how her impact has helped me achieve this accomplishment; I am very excited to be able to share this with Bianca.

I also want to thank *another* Dr. Baron, Dr. Barbara Baron, who I volunteered with during high school at an organization for youth with Tourette's Syndrome. You inspired me to become a psychologist, so without you I likely would not be here today.

While both Dr. Barons and Dr. Cody Murphy sparked my clinical interests, the Lab for Computational Neuroimaging at Mass General sparked my interest in neuroimaging and provided me with a research foundation to succeed in grad school. I am especially thankful for Allison Stevens and André van der Kouwe, who I look up to immensely. You have been incredible role models in both work and life.

I am extremely grateful to my parents, Marc and Kathy, for always supporting my academic dreams and doing whatever they could to help me accomplish them. I also want to thank my sisters, Andrea and Ashley, for always being there for me even while living far apart in different parts of the country. My aunts and uncles have also been a wonderful source of encouragement since I was a young girl.

Finally, I want to give enormous thanks to my partner, Justin Brooks, who has been the most incredible support throughout this very challenging and wild ride of graduate school. You helped me keep going with your encouragement, hugs, and delicious meals on nights when I was working away, on top of reading drafts, reviewing abstracts, and polishing figures, all while making me laugh when I needed it most.

Allison Moreau

Washington University in St. Louis

August 2023

ABSTRACT OF THE DISSERTATION

Brain Structure Correlates of Obsessive-Compulsive Personality Traits

by

Allison Moreau

Doctor of Philosophy in Psychological and Brain Sciences

Washington University in St. Louis, 2023

Professor Ryan Bogdan, Chair

Obsessive-Compulsive Personality Disorder (OCPD) is the most common personality disorder, yet much remains unknown about its etiology. Although neural contributions to many other psychiatric disorders have been extensively studied, few existing studies have examined neural correlates of OCPD. Furthermore, all have had insufficient sample sizes to produce reliable results. Large samples are needed to reliably detect the expected small brain-behavior relationships. However, large neuroimaging studies often do not assess for personality disorders, although many assess for normative personality. The present study employed a Five-Factor Model of personality disorders, which conceptualizes personality disorders as maladaptive extremes of normative personality traits, to predict OCPD scores from normative personality data using machine learning techniques in a large community-based sample (n=1,606). This trained ML model was then applied to a separate dataset with normative personality and neuroimaging data (n=1,253) to generate predicted OCPD scores and subsequently examine brain structure correlates of OCPD traits. Despite a moderate ability to predict OCPD traits using normative personality data that generalizes across samples, we found limited evidence that predicted OCPD scores are associated with individual differences in brain structure. Indeed, there was only one significant univariate association wherein thicker right superior frontal cortex was associated with higher OCPD scores.

Adopting ML models to generate multivariate models of brain structure resulted in imprecise models and thus no reliable associations. Collectively, these data suggest that OCPD symptoms may be predicted using normative personality data, but that OCPD personality traits may not be strongly associated with brain structure and may require exceptionally large samples to reliably identify these modest associations. Broadly, this approach exemplifies how deeply phenotyped small samples may be used to inform large national samples that may not have assessed specific phenotypes.

Chapter 1: Introduction

Obsessive-Compulsive Personality Disorder (OCPD) is the most common personality disorder, with a prevalence rate of 4.3% in Western countries (Volkert et al., 2018). It is characterized by "a pervasive pattern of preoccupation with orderliness, perfectionism, and mental and interpersonal control, at the expense of flexibility, openness, and efficiency" (American Psychiatric Association, 2013, p. 678) that results in occupational and/or social impairment (e.g., not finishing tasks and missing deadlines due to perfectionism and refusing to delegate or alienating friends and family with one's rigidity and insistence on control) (Skodol et al., 2002, 2005). This occupational stress and impairment has been linked to burnout, depression, and physical health conditions (Atroszko et al., 2020). Some evidence suggests that executive functioning is also impaired (García-Villamisar & Dattilo, 2015). In addition to these negative impacts on individuals, OCPD also presents an economic burden, estimated to be over $1,400 per person annually in direct and indirect costs (Soeteman et al., 2008). The individual and societal costs of OCPD make it important to understand its etiology. As individual differences in behavior have been associated with variability in brain structure and function, studying the neural correlates of OCPD may help inform our etiologic understanding of this debilitating disorder as well as refine its nosological distinctions (Morris & Cuthbert, 2012).

1.1 Existing OCPD Neuroimaging Literature

Unlike most other psychiatric disorders, there have been few neuroimaging studies of OCPD; indeed, I am only aware of seven published studies (**Table 1.1**). Four of the publications evaluated brain structure; of these four, three are from the same small sample (total n=36 [50% patients])

and the fourth did not examine OCPD independently. The remaining three independent published studies evaluated brain function.

1.1.1 Brain Structure Studies

Structural neuroimaging studies can examine a variety of metrics including the volume, surface area, or thickness of brain regions, as well as properties of structural connectivity that characterize the white matter fiber tracts that connect different regions of the brain. Structural connectivity is measured by a type of magnetic resonance imaging (MRI) called diffusion weighted imaging (DWI); diffusion tensor models can be "fit" to these images to measure the orientation and water diffusion properties of the fiber tracts. Common structural connectivity metrics include fractional anisotropy (FA) (the directionality of diffusion), mean diffusivity (MD) (the average rate of diffusion), axial diffusivity (AD) (the rate of diffusion along the main axis of the tract) and radial diffusivity (RD) (the rate of diffusion in the transverse direction of the tract) (Soares et al., 2013). Existing OCPD studies have examined all of these metrics.

The three OCPD structural publications from the same sample (N=16 patients, 18 controls) (Atmaca, Korucu, Caglar Kilic, et al., 2019; Atmaca, Korucu, Tabara, et al., 2019; Gurok et al., 2019) reported that participants with OCPD had smaller volumes of the orbitofrontal cortex, amygdala, hippocampus, and pineal gland and larger thalamic volume. The fourth structural study (Payer et al., 2015) reported that Cluster C personality disorder symptoms, which includes OCPD, were associated with larger caudate tail surface area, smaller ventral striatum volume, and thicker right prefrontal cortex. While the study did not analyze OCPD symptoms independently, 25 of the 29 participants in the Cluster C group had elevated OCPD symptoms, highlighting their relevance for understanding OCPD. Finally, an unpublished dissertation examined structural connectivity in OCPD (Fernandes Gonçalves, 2015). In a sample of 18 participants, that may overlap with one of

the functional studies, individuals with OCPD (N=9) had higher mean diffusivity (MD) in the right cingulum bundle and higher axial diffusivity (AD) in the tract connecting the right precuneus and lateral occipital cortex.

Several limitations strongly reduce the informativeness of these prior brain structure studies of OCPD. First, the sample sizes were small (ns=18-72), making it likely that effect sizes are overestimated and possible that the findings are false positives (Button et al., 2013). Given expected small effects between complex behavioral traits and brain phenotypes (Dick et al., 2020; Marek et al., 2022), as well as the heterogeneity of OCPD, these studies are likely substantially underpowered. Second, the published data are relatively lower resolution (1.5T field strength, 2.4 mm slice thickness) than current standards in the field which results in less accurate image segmentation and resulting estimates of brain structure. Third, the three publications from the same sample evaluated independent regions of interest; if multiple testing was adjusted for across these studies, some results would not withstand multiple testing correction (12 tests across sample; Bonferroni corrected alpha level = 0.0042; thalamus and OFC findings no longer significant). Finally, it is unclear if the Atmaca studies controlled for age, sex, and intracranial volume simultaneously.

1.1.2 Brain Function Studies

Three studies have examined brain function in OCPD, including two resting state fMRI studies and one electroencephalography (EEG) study. Coutinho and colleagues (2016) reported increased resting-state functional connectivity in the left precuneus of OCPD patients in a small (N=20; 50% OCPD) pilot study. Meanwhile, Lei and collaborators (2020) (N=74, 50% OCPD) found that individuals with OCPD had higher amplitudes of low frequency fluctuation (ALFF), a measure of spontaneous neural activity at rest, in the bilateral caudate and left precuneus, insula, and medial

superior frontal gyrus, as well as lower ALFF in the left lingual and right fusiform gyri. Left precuneus ALFF values correlated with OCPD severity. The EEG study (Luo et al., 2020) compared neural mechanisms of OCPD and obsessive-compulsive disorder (OCD) during decision making and found that OCPD and OCD subjects (Ns=19, 24, respectively) both had larger feedback-related negativity waveforms than healthy controls (N=26); feedback-related negativity is an EEG signal of brain activity that is sensitive to valenced feedback (Rawls et al., 2020).

Table 1.1 Existing OCPD Neuroimaging Studies

Study	Imaging	Sample	Findings
Structural			
Atmaca et al. (2019a), Atmaca et al. (2019b), Gurok et al. (2019)	Volume *(1.5T)*	34 Total 16 OCPD 18 Controls	OCPD subjs had smaller volume in OFC, amygdala, hippocampus and pineal gland, larger thalamus volume. No global volume differences.
Payer et al. (2015)	Volume, cortical thickness, and surface area *(1.5T)*	72 Total 37 w/ likely PD (20 Cluster B, 28 Cluster C, 11 Cluster B + C, 0 Cluster A) 35 Controls	Cluster C PD-sxs associated w/ larger caudate tail surface area, smaller ventral striatum volume, thicker cortex in RH superior and middle frontal gyri, lateral OFC
Fernandes Gonçalves (2015)* *unpublished*	Diffusion *(3T)*	18 Total 9 OCPD 9 Controls	OCPD subjs had higher MD in RH cingulum bundle + higher AD in precuneus to lateral occipital cortex tract
Functional			
Coutinho et al. (2016)	Resting-state fMRI *(3T)*	20 Total 10 OCPD 10 Controls	OCPD subjs had increased resting-state functional connectivity in LH precuneus (posterior DMN node)
Lei et al. (2020)	Resting-state fMRI *(3T)*	74 Total 37 OCPD 37 Controls	OCPD subjs had higher ALFF in bilateral caudate, LH precuneus, LH insula, LH medial superior frontal gyrus + lower ALFF in RH

			fusiform and LH lingual gyri. LH precuneus ALFF values correlated w/ OCPD severity scores
Luo et al. (2020)	EEG	69 Total 19 OCPD, 24 OCD 26 Controls	OCD + OCPD subjs had larger feedback-related negativity (FRN) waveform (lose-win) than controls

Notes: ALFF = amplitude of low frequency, AD = axial diffusivity, DMN = default mode network, EEG = electroencephalography, fMRI = functional MRI, LH = left, MD = mean diffusivity, OCD = obsessive compulsive disorder, OCPD = obsessive compulsive personality disorder, OFC = orbitofrontal cortex, PD-sxs = personality disorder symptoms, RH = right, subjs = study participants.

1.1.3 Summary of Existing Studies

While the literature contains no replicated findings, several brain regions (i.e., precuneus, caudate, prefrontal cortex) were linked to OCPD across multiple studies. The precuneus, a node of the default mode network, is associated with self-referential processing and rumination (Cavanna & Trimble, 2006; Zhou et al., 2020). The default mode network is active when the brain is "at rest" (Raichle, 2015). Individuals with OCPD often experience high levels of preoccupation and a ruministic cognitive style (Smith et al., 2006), making precuneus findings salient. The left precuneus exhibited increased resting state functional connectivity (Coutinho et al., 2016) and ALFF (Lei et al., 2020) while a white matter tract connecting the right precuneus and lateral occipital cortex exhibited increased axial diffusivity (Fernandes Gonçalves, 2015). The caudate, which is involved in executive planning and goal-directed action (Grahn et al., 2008), showed increased surface area of its tail (Payer et al., 2015) and ALFF (Lei et al., 2020) in both hemispheres. In OCPD, a hyperfocus on goal-direction action, often at the expense of interpersonal relationships and efficiency, is common. The prefrontal cortex coordinates most higher-level cognition, including cognitive control, which covers a range of mental processes such as inhibiting

automatic responses, monitoring and planning, and shifting between tasks (N. P. Friedman & Robbins, 2022). Individuals with OCPD often exhibit behavioral differences in these processes, including a strong ability to resist impulsive urges and other automatic responses and excessive planning. In OCPD, the prefrontal cortex exhibited increased ALFF in the left medial superior frontal gyrus (Lei et al., 2020) and thicker cortex in the right superior and middle frontal gyri (Payer et al., 2015). Studies also found both smaller volume in the bilateral orbitofrontal cortex (Atmaca, Korucu, Tabara, et al., 2019) and thicker cortex in right lateral orbitofrontal cortex (Payer et al., 2015). This sparse and inconsistent literature suggests that more research is needed before a model of OCPD neurobiology can be advanced (Marincowitz et al., 2021).

Although some patterns across imaging modalities have emerged from the existing literature, all of the studies have been limited by small sample sizes (Button et al., 2013). Large samples are needed for well-powered neuroimaging studies capable of detecting the small effects often found for brain-behavior relationships (Dick et al., 2020). However, most large neuroimaging studies do not explicitly assess for personality disorders. Therefore, an alternative method of assessing personality disorders is needed for a sufficiently powered neuroimaging study of OCPD-related traits. The five-factor model of personality, which is readily assessed in many large consortium and independent neuroimaging samples, provides this alternative method and has proved useful for other personality disorders (e.g., (Baranger et al., 2020)).

1.2 Five Factor Model of Personality Disorders

A popular dimensional model of personality psychopathology argues that personality disorders represent maladaptive extremes of personality traits (Widiger & Trull, 2007). The most prevalent model of normative personality traits is the Five Factor Model (FFM), which includes the traits

neuroticism, extraversion, openness, agreeableness, and conscientiousness (Digman, 1990). Neuroticism describes the tendency to experience distress or negative affect and subsequent cognitive and behavioral patterns. Extraversion represents the tendency to obtain gratification from the outer world rather than inner self. Openness describes a tendency to appreciate and seek out novel ideas and experiences. Agreeableness represents a tendency to prioritize social harmony in interpersonal interactions. Conscientiousness describes a tendency to exhibit self-discipline and achievement-striving. Each trait has been further broken down into more specific facets (e.g., the extraversion facet assertiveness). The five-factor model of personality disorders characterizes the disorders by the extent to which higher-level traits (or "domains") and lower-level facets of the five-factor model are exhibited in the disorders.

Meta-analyses broadly support the FFM-based conceptualization of personality disorders (Samuel & Widiger, 2008; Saulsman & Page, 2004). More specifically, using data from 2,873 participants from 12 studies Saulsman and Page (2004) found that the five-factor model was meaningfully associated with all 10 personality disorders. Most personality disorders were positively associated with neuroticism and negatively associated with agreeableness. The correlations between the five personality domains and the personality disorders ranged from |0.01 to 0.49|. A subsequent meta-analysis that extended Saulsman and Page's work by examining the 30 facets of the FFM, in addition to the "Big Five" traits or "domains", also supported the FFM conceptualization of personality disorders (Samuel & Widiger, 2008). When combining results from 3,207 individuals across 16 studies and 18 independent samples (only one of which was in Saulsman and Page (2004)), each personality disorder had significant correlations between researchers' and clinicians' hypothesized FFM profiles and meta-analytic results, with all but one (histrionic PD) greater than 0.50.

This support for a dimensional model of personality disorders led to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) including a proposed dimensional model in its "Emerging Measures and Models" section (American Psychiatric Association, 2013). While clinical practice still uses the same categorical approach and diagnostic criteria as the DSM-IV, the proposed model involves a dimensional assessment of personality functioning and pathological personality traits that are maladaptive variants of the FFM domains (although one domain, psychoticism, appears less directly linked to the "Big Five"). The five maladaptive personality domains are: negative affectivity (i.e., neuroticism), detachment (the maladaptive opposite of extraversion), antagonism (the maladaptive opposite of agreeableness), disinhibition (the maladaptive opposite of conscientiousness), and psychoticism. Similar to the FFM, each of these domains is broken down into lower-order facets, totaling 25 in all. Under this alternative model for personality disorders (AMPD), OCPD is diagnosed when at least three out of four pathological personality traits (i.e., rigid perfectionism (required), perseveration, intimacy avoidance, and restricted affectivity) are present and associated with impairment in personality functioning.

1.2.1 FFMs of OCPD

While overall the FFM model of personality disorders has received strong support, the FFM conceptualization of OCPD specifically has been controversial. Clinicians and researchers have proposed that OCPD can be represented by high levels of anxiousness and all of the conscientiousness facets along with low levels of excitement seeking, openness to actions, and openness to values (Lynam & Widiger, 2001; Samuel & Widiger, 2004). Researchers have also suggested that low levels of impulsiveness and openness to feelings and ideas are present (Lynam & Widiger, 2001). However, despite these hypothesized representations and evidence that most

personality disorders show high correlations with normative personality traits, the empirical data for OCPD have been less strong.

While personality disorders tend to have low convergent validity broadly (r=0.60 for two self-report measures and r=0.39 for self-report and clinician interview) (Widiger & Boyd, 2009), meta-analysis has revealed that the convergent validity between FFM OCPD scores and DSM-IV OCPD symptom counts is lower (r=0.16) than any other personality disorder (Miller, 2012). This can also be seen in low correlations between traditional OCPD assessment scores with both FFM domains (|r|=|0.04-0.24|) and facets (|r|=|0.01-0.25|) in meta-analyses (**Table 1.2**) (Samuel & Widiger, 2008; Saulsman & Page, 2004). For both domains and facets, conscientiousness exhibited the largest effect sizes. (Samuel & Widiger, 2008). For 22 of the 30 facet effect sizes, there was also evidence of between-study heterogeneity in results. The DSM-5 AMPD model of OCPD, which was established after these meta-analyses, has also not performed well. Studies have found that the DSM-5 AMPD traits account for only 24 to 38% of the variance in OCPD scores (reviewed in (Samuel et al., 2022)). Collectively, meta-analyses suggest that FFMs of OCPD do not well represent prior definitions of OCPD.

However, the meta-analytic OCPD profile created by facet-level results correlates strongly with consensus FFM OCPD profiles from researchers (r=0.92 (Lynam & Widiger, 2001)) and clinicians (r=0.91 (Samuel & Widiger, 2004)). Thus, these data suggest that while FFMs of OCPD may not well represent prior OCPD diagnostic classification, they align well with researcher and clinician conceptualizations of the disorder. This overlap between FFM meta-analytic results and researcher and clinician descriptions of OCPD suggests that an FFM model of OCPD holds potential. Improvements to FFM measures of OCPD are needed to realize that potential.

1.2.2 Improving FFM OCPD Measurement

Incorporating Maladaptive Conscientiousness

The unexpectedly small correlations between conscientiousness and OCPD may be due to normative FFM personality assessments not capturing maladaptive levels of conscientiousness. For example, in the revised NEO-Personality Inventory (NEO-PI-R), the most commonly used FFM measure, only 10% of high conscientiousness items describe maladaptive or dysfunctional behavior (Haigler & Widiger, 2001). After modifying items to describe more extreme or maladaptive versions of the trait, Haigler and Widiger found that the correlations between conscientiousness and OCPD were much higher, 0.47-0.69 compared to -0.02-0.27 in the original NEO-PI-R. Samuel and Widiger (2011) administered multiple measures of conscientiousness, OCPD, and specific components of OCPD to college students (n=536 oversampled for OCPD). They found that conscientiousness instruments designed to assess general personality functioning had small to medium relationships with OCPD, but instruments designed for maladaptive versions of traits had large effect sizes. These findings suggest that FFMs of normative personality traits may contribute noise to OCPD by being unable to distinguish between adaptive and maladaptive conscientiousness.

Heterogeneity in OCPD Assessment

Some data suggests that heterogeneity in OCPD assessments, as opposed to FFMs of OCPD, may weaken correlations between conscientiousness and OCPD. For instance, a meta-analysis comparing correlations between FFM facets and OCPD for several different FFM and PD measures found that two of the PD measures, the Millon Clinical Multiaxial Inventory (MCMI-III) and Schedule for Nonadaptive and Adaptive Personality (SNAP), had relatively larger correlations with conscientiousness facets (MCMI-III: 0.38-0.5, SNAP: 0.17-0.3) while two others, the PDQ and SCID-II, produced lower estimates (all rs $\leq$0.09) (Samuel & Widiger, 2008).

In a follow-up study comparing eight OCPD measures, Samuel and Widiger (2010) reported that the OCPD measures only had moderate convergent validity, with important differences in how they captured conscientiousness, agreeableness, and neuroticism. The correlations between these personality traits and OCPD differed in direction (i.e., positive or negative) and magnitude (e.g., conscientiousness r's from -0.06 to 0.71) across the OCPD measures. Clearly, the OCPD assessment tool used strongly impacts how well an FFM model of OCPD performs.

Heterogeneity in Personality Factors

Researchers have also investigated if different statistical approaches for measuring the relationship between conscientiousness and OCPD might also demonstrate stronger correlations between the two. Mike and colleagues (2018) used a bifactor model to distinguish common trait variance from facet-specific variance (**Table 1.2**). Surprisingly, they found that conscientiousness was negatively associated with OCPD (r=-0.22). At the facet level, however, achievement striving, order, and dutifulness were all positively and significantly associated with OCPD (*r*s=0.55, 0.39, and 0.25, respectively). Notably these bifactor-generated correlations are much higher than those found in studies using zero-order correlations. The authors suggest that OCPD is related to higher levels of these facets, individually, yet when combining all of these personality traits into the domain of conscientiousness, the unexpected negative correlation between conscientiousness and OCPD may be due to the perfectionism and rigidity of OCPD hindering individuals' abilities to accomplish tasks and goals, a core element of conscientiousness. These results suggest that facet-level analyses are needed and that looking at conscientiousness as a single factor may introduce heterogeneity that attenuates correlations with OCPD.

Table 1.2. Correlations between FFM and OCPD

	Saulsman and Page (2004)	Samuel and Widiger (2008)	Mike et al. (2018)	Mike et al. (2018)
Neuroticism	**0.08**	**0.18**	**0.45**	**0.20**
Anxiousness		**0.16**	0.07	**0.18**
Angry hostility		**0.10**	**0.33**	**0.11**
Depressiveness		**0.09**	-----	-----
Self-consciousness		**0.13**	0.08	**0.12**
Impulsiveness		**−0.07**	-0.02	0.05
Vulnerability		**0.03**	-0.06	0.05
Extraversion	**-0.12**	**-0.12**	**-0.27**	**-0.11**
Warmth		**−0.07**	-0.04	0.04
Gregariousness		**−0.16**	**-0.28**	**-0.32**
Assertiveness		**−0.01**	0.04	0.02
Activity		**0.03**	**0.12**	**0.11**
Excitement seeking		**−0.12**	**0.24**	0.06
Positive emotions		**−0.09**	-0.10	**-0.10**
Openness	**-0.07**	**-0.04**	-0.10	-0.02
Fantasy		**−0.09**	**0.10**	0.02
Aesthetics		**0.01**	-0.08	0.06
Feelings		**0.01**	**0.16**	0.11
Actions		**−0.12**	**-0.44**	**-0.35**
Ideas		**0.03**	0.06	0.13
Values		**−0.09**	**-0.21**	-0.12
Agreeableness	-0.04	**-0.05**	**-0.50**	**-0.19**
Trust		**−0.08**	**-0.17**	-0.02
Straightforwardness		**0.04**	0.09	0.06
Altruism		**0.04**	-----	-----
Compliance		**0.01**	-0.13	-0.01
Modesty		**0.02**	0.02	0.08
Tendermindedness		**0.00**	-----	-----
Conscientiousness	**0.23**	**0.24**	**-0.22**	**-0.17**
Competence		**0.19**	0.05	0.09
Order		**0.25**	**0.39**	**0.27**
Dutifulness		**0.25**	**0.25**	**0.23**
Achievement striving		**0.25**	**0.55**	**0.55**
Self-discipline		**0.21**	-----	-----
Deliberation		**0.24**	0.05	0.07
Data source information	Meta-analysis	Meta-analysis	NEO & self-report MAPP	NEO & interviewer SIDP
Effect Statistic	Pearson's correlations	Pearson's correlations	Bi-factor correlations	Bi-factor correlations
Significance	Bold values significant at $p<0.001$ (one-tailed)	Bold values significant at $p<0.05$	Bold values indicate 95% CI does not include 0	Bold values indicate 95% CI does not include 0

Note: CI = confidence interval, MAPP = Multi-source Assessment of Personality Pathology, NEO = Revised NEO Personality Inventory (NEO-PI-R), SIDP = Structured Interview for DSM-IV Personality. Cells with dashed lines for Mike et al. (2018) indicate facets that did not have significant variance after accounting for the general trait. Saulsman and Page (2004) only examined personality domains, so facet level data are unavailable.

Five-Factor Obsessive-Compulsive Inventory (FFOCI)

These findings on maladaptive conscientiousness and the heterogeneity of measures, along with others (Clark et al., 1996; Markon et al., 2005; O'Connor, 2005; Schroeder et al., 1992; Watson et al., 2008) in support of the FFM model of OCPD, led Samuel and colleagues to develop a five-factor measure of obsessive-compulsive personality traits, the Five-Factor Obsessive-Compulsive Inventory (FFOCI) (Samuel et al., 2012). FFOCI items were written to capture maladaptive variants of the relevant FFM facets (competence ("perfectionism"), order ("fastidiousness"), dutifulness ("punctiliousness"), achievement striving ("workaholism"), self-discipline ("doggedness"), deliberation ("ruminative deliberation"), warmth ("detached coldness"), excitement-seeking ("risk aversion"), openness to feelings ("constricted"), actions ("inflexibility"), values ("dogmatism"), and anxiety ("excessive worry")). The FFOCI demonstrated significant convergent validity with four existing measures of OCPD (convergent correlations ranged from 0.50 to 0.71). While the FFOCI also provided significant incremental validity over the NEO-PI-R facets for predicting an OCPD composite (ΔR^2 from 0.03 to 0.39, mean 0.2), it is important to note that all but one of the 12 NEO-PI-R facets predicted a significant portion of variance in an OCPD composite (Samuel et al., 2012). This suggests that NEO-PI-R-based measures may still be valid in samples where specialized measures such as the FFOCI are unavailable.

1.2.3 Summary of FFM Models of Personality Disorders

In summary, meta-analytic research broadly supports FFM models of personality disorders (Samuel & Widiger, 2008; Saulsman & Page, 2004), which led to its inclusion as an alternative characterization of PDs in the DSM-5. However, among PDs, OCPD was least well captured by normative FFMs of personality despite aligning well with researcher and clinician versions of the disorder. Research suggests this is likely due to commonly used FFM measures not capturing maladaptive conscientiousness as well as the variability in OCPD measures, rather than conceptual flaws with a FFM model of OCPD. The development of the FFOCI provides a FFM measure of OCPD that addresses these concerns.

1.3 FFM Neuroimaging Studies

To my knowledge only two existing studies (Baranger et al., 2020; Haas & Miller, 2015) utilize a FFM model of personality disorders to examine neural correlates. However, the field of personality neuroscience contains numerous neuroimaging studies examining the neural correlates of the five-factor model. Although both functional and structural neuroimaging studies have examined personality and personality disorders, the current study focuses on brain structure given increasing evidence that task-based fMRI studies may not be sufficiently reliable for individual differences research (Elliott et al., 2020) and resting-state fMRI acquisitions must be much longer than originally thought to obtain reliable data (Gordon et al., 2017).

The normative personality literature has generated mixed and inconsistent findings regarding brain structure. Several studies analyzing data from the Human Connectome Project (HCP) (Hyatt et al., 2019; Owens et al., 2019; Riccelli et al., 2017) as well as other, smaller, samples (Bjørnebekk et al., 2013; DeYoung et al., 2010; Privado et al., 2017; Vartanian et al.,

2018) have found significant associations between each of the "Big Five" and brain structure. However, other studies (Avinun et al., 2020; Masouleh et al., 2019), including some also analyzing HCP data (Gray et al., 2018; Nostro et al., 2017; Valk et al., 2020), have found less or no support for brain structure correlates of personality. Hyatt and colleagues examined whether these null to small effect sizes were due to the level of personality and/or morphometry being analyzed (Hyatt et al., 2022). Using mixed effects models in the HCP dataset, higher-order variables (domains for personality and "omnibus" measures for morphometry (e.g., total brain volume, total cortical area, mean cortical thickness)) had the largest effect sizes on average, although these relationships were still small. Together, these findings suggest that personality traits may not be robustly associated with brain structure as measured by typical univariate analyses, or, if they are, large samples are needed to accurately estimate effects.

1.4 The Neural Correlates of OCPD: Gaps in the Literature

In summary, the existing neuroimaging studies of OCPD have been too small to provide reliable results. Larger samples are required for future studies; however, it is difficult to obtain large neuroimaging samples that include explicit personality disorder assessments. This leads to the need for another way of measuring personality disorders, which the five-factor model of personality disorders offers, given its basis in a normative personality taxonomy that is more frequently assessed in large studies. Yet, the controversy over the FFM conceptualization of OCPD and the lack of robust, replicable personality neuroimaging findings present challenges to using the typical approaches to measuring and analyzing FFM OCPD traits (i.e., count or similarity scores) and neuroimaging data (i.e., numerous regressions of individual brain regions). To overcome these limitations, multivariate machine learning is a promising alternative to traditional factor analytic

approaches to determine an FFM-based model of OCPD. In addition, it may also be better able to identify the small effect sizes expected in the neuroimaging results by aggregating large amounts of data (Abi-Dargham & Horga, 2016; Reddan et al., 2017; Walter et al., 2019a).

1.5 Machine Learning

Machine learning is a branch of computer science that uses algorithms to learn to recognize patterns and make predictions from data. This can provide information about the data at hand or allow predictions to be made about new data. Machine learning (ML) approaches are typically divided into two main types: supervised and unsupervised. Supervised ML techniques use labeled training data to predict a categorical (i.e., classification) or continuous (i.e., regression) variable in novel test data. Unsupervised ML techniques detect patterns from unlabeled training data which can then be characterized and applied to new datasets. These unsupervised approaches often focus more on grouping observations rather than labeling them. This offers an alternative to expert-perceived similarities in symptoms to drive diagnoses and allows for data to generate observed clusters that researchers and clinicians may not detect.

While classical statistical methods test for group differences, ML approaches are typically used for prediction, which in psychiatry often means single-subject diagnostic or prognostic prediction. For example, machine learning has been used to predict treatment response (Webb et al., 2020), to differentiate individuals with anxiety versus depression (Richter et al., 2020), to identify the most informative risk factors for a disorder (Beeney et al., 2021), and even to predict suicidal behavior without data on suicidal ideation (Horvath et al., 2020). Thus, machine learning holds potential for a precision medicine approach in psychiatry; this potential is a large reason for the markable growth of machine learning in psychiatric research (Shatte et al., 2019).

Machine learning techniques have also increased in popularity due to several additional advantages. Machine learning approaches can model both linear and non-linear relationships, account for variable interactions, and use both correlations between and uniqueness of variables to predict outcomes. They also reduce the variance of predictions, providing more stability across samples. Thus, machine learning provides a more robust way to model disorders and associated clinical features.

For personality disorders, machine learning may be able to identify how the interplay among normative personality traits are associated with the disorders. Given the large number of traits and underlying facets identified in dimensional models of personality disorders, they are well-positioned to benefit from ML approaches which adeptly handle high numbers of features. Using a supervised machine learning approach in combination with an epidemiological sample that includes explicit measures of personality disorders holds potential for a more refined conceptualization of personality traits associated with personality disorders.

In neuroimaging data, machine learning's ability to consider hundreds or thousands of variables (or "features") at once permits researchers to look for combinations of neural metrics (e.g., brain region volumes, BOLD signal levels during tasks) most associated with outcomes. Given the small effect sizes typically seen in standard neuroimaging analyses that test individual voxels or regions of interest separately, identifying a combination of neural phenotypes that are correlated with outcomes of interest may result in stronger, more clinically meaningful effects (Abi-Dargham & Horga, 2016; Reddan et al., 2017; Walter et al., 2019a). Understanding the relationships between identified phenotypes may also lead to new mechanistic understanding of the disorder or phenomenon of interest.

1.6 Aims and Hypotheses

The aims of the current study are to 1) develop and validate a predictive model of OCPD scores generated from normative personality data, and 2) apply this model to a large neuroimaging dataset that does not have explicit OCPD assessment data to examine brain structure correlates of OCPD traits. I hypothesize that OCPD traits will be associated with structural differences in the prefrontal cortex, including the orbitofrontal, dorsal and rostral anterior cingulate cortices, medial and lateral prefrontal cortices, as well as the insula, posterior cingulate cortex, precuneus, caudate and ventral striatum. The orbitofrontal cortex is involved in decision making and subjective valuation (Bechara A. et al., 2000; Walton et al., 2011), which may be altered in individuals with OCPD such that tasks and rule-following are valued over leisure activities and flexibility. The dorsal anterior cingulate cortex (dACC) plays an important role in behavioral control and conflict monitoring (Carter & van Veen, 2007; Shenhav et al., 2016), which can both be conceptualized as overexpressed in OCPD. Differences in the dACC may make it harder for individuals with OCPD to correct their actions based on differences between expectations and actual outcomes which precludes their ability to be effective. The rostral anterior cingulate cortex is involved in cognitive control and serves as a hub where motivation and action control networks interact (Tang et al., 2019). Alterations in this region could be linked to the rigid over-control present in OCPD. The dorsolateral PFC (dlPFC) is responsible for forming plans (Kaller et al., 2011; Mushiake et al., 2006), so impairments in this region could lead to the perseveration on lists and plans that often get in the way of finishing tasks. The right inferior frontal gyrus acts as a brake for response tendencies (Aron et al., 2014). Differences in this region could explain individuals with OCPD's strong ability to inhibit impulsive tendencies. The insula controls interoception to both sensorimotor and socioemotional cues and assigns subjective value to body signals that it sends to

the prefrontal cortex (Mutschler et al., 2009). It has also been linked to empathy and flexible behavior (Singer et al., 2009); thus, it could be involved in the scrupulousness and rigidity of OCPD. The posterior cingulate cortex and precuneus, two key nodes of the default mode network, are both involved in self-referential thought and rumination (Brewer et al., 2013; Cavanna & Trimble, 2006; Zhou et al., 2020), especially for future events, so differences in these regions may correspond to the perfectionism and preoccupation with plans and schedules that predominate in OCPD. The caudate is involved in executive planning and response switching (Grahn et al., 2008), through its connections with the dlPFC, while the ventral striatum is involved in motivation (Cardinal et al., 2002), so both of these subcortical regions may also exhibit impairments in individuals with OCPD.

Chapter 2: Methods

2.1 Studies

Three independent study samples were used in the present study (total n=3,034), the: 1) St. Louis Personality and Aging Network (SPAN) Study (Oltmanns & Gleason, 2011); 2) Five Factor Obsessive Compulsive Inventory (FFOCI) Validation study (Samuel et al., 2012); 3) Duke Neurogenetics Study (DNS) (Nikolova et al., 2012). The SPAN study is an ongoing study of personality, experiences, biology, and health. It was leveraged to generate a predictive model of OCPD using normative personality data. The FFOCI validation study was conducted to investigate the psychometric properties of the FFOCI. It was used to test the external validity of the predictive OCPD model in an independent sample of young adults. The DNS study examined a wide range of psychological constructs, behavioral phenotypes, genetics, and the brain to explore their interactions in health and psychopathology. It was used to examine associations between predicted OCPD scores and brain structure. All studies followed protocols approved by relevant institutional review boards.

2.2 Participants

2.2.1 OCPD Model Training Sample – St. Louis Personality and Aging Network (SPAN) Study

The St. Louis Personality and Aging Network (SPAN) study is an ongoing longitudinal protocol that recruited 1,630 mid-late life adults (55-64 years of age) for the initial baseline session from 2007-2011 in the St. Louis area (**Table 2.1**). Participants were recruited using listed phone numbers that were crossed with census data to identify households with at least one member in the

target age range. When more than one person was in the target age range, the Kisch (1949) method was used to randomly select one individual from the household. If the target refused to participate, other potentially eligible residents were not recruited. Individuals were excluded if they lacked permanent residence, could not read at a 6th-grade level, or had active psychotic symptoms. The SPAN sample is 55% female, and 66% Caucasian, 32% African American, and 2% another reported race/ethnicity. 1.9% of participants identified as Latinx. Regarding education level, 1.7% did not complete high school, 32.1% received a high school diploma or equivalent, 14.2% completed a 2-year college degree or vocational school, 25.2% had a bachelor's degree, and 26.9% had an advanced degree. At baseline, participants completed a 3-hour in-person interview and self-report questionnaires, followed by brief mailed/online follow-up questionnaires every 6 months. In-person follow-up interviews occurred every 2.5-3.5 years after baseline. Data from the baseline (n=1,630) and two in-person follow-up visits (FU10, n=1,344; and FU12, n=937) were considered. The analytic sample sizes, after addressing missing data (described below in section 2.4.1), were 1,606 (baseline), 1,015 (FU10), and 898 (FU12).

2.2.2 OCPD Model External Validation Sample – FFOCI Validation Study

The FFOCI Validation study was used to test the external validity of the trained OCPD model and its extension to a college-aged sample (**Table 2.1**) (Samuel et al., 2012). Data from 203 college students included in the study's psychometric validation analyses were considered. The sample (n=203) is 64.9% female, and 88% Caucasian, 4.5% African American, 3.5% Asian, 2.0% multiracial, and 2.0% another reported race. 1.0% of participants identified as Latinx. The average age was 19.4 (SD = 4.5), with most participants falling between 18 to 26 years old and two older (i.e., 46 and 71 years old) students. The analytic sample size after addressing missing data (described below in section 2.4.2) was 175.

2.2.3 Neuroimaging Sample – Duke Neurogenetics Study (DNS)

The Duke Neurogenetics Study (DNS) assessed a wide range of behavioral and biological traits among 1,330 young adult college students (**Table 2.1**). Study exclusions included: major medical diagnoses (e.g., cancer), lifetime history of psychotic symptoms, use of psychotropic or hypolipidemic medications, and conditions affecting cerebral blood flow (e.g., hypertension). The sample is 57% female, and 49.9% Caucasian, 27.3% Asian, 11.8% African American/Black, 7.8% multiracial, 0.2% Native American, and 2.9% another reported race. 10.2% of participants identified as Latinx. The average age was 19.7 (SD =1.3). Participants completed two sessions between 2012-2016: 1) a neuroimaging session, and 2) a behavioral assessment and clinical interview. The current sample (n=1,253) is comprised of participants with acceptable neuroimaging and complete personality data.

Table 2.1. Participant Demographics

Variable	SPAN (N=1,606)	FFOCI Validation Study (N=175)	DNS (N=1,253)
Age	Range: 55-64 Mean (SD): 60.0 (2.7)	Range: 18-71 Mean (SD): 19.4 (4.6)	Range: 18-22 Mean (SD): 19.7 (1.3)
Gender	884 female (55.0%)	115 female (65.7%)	727 female (58.0%)
Race			
Caucasian/White	1,052 (65.7%)	152 (87%)	628 (50.1%)
Asian	5 (0.3%)	7 (4.0%)	337 (26.9%)
African American/Black	513 (32.0%)	8 (4.6%)	149 (11.9%)
Native American	3 (0.2%)	0 (0%)	2 (0.2%)
Multiracial	9 (0.6%)	3 (1.7%)	101 (8.1%)
Other	19 (1.2%)	4 (2.3%)	36 (2.9%)
Ethnicity – Hispanic/Latino	30 (1.9%)	1 (0.5%)	126 (10.1%)

Note: Summary statistics are provided for the subjects that will be included in analyses based on complete data. For SPAN and the FFOCI validation study, this includes subjects with NEO and OCPD data. For DNS, this includes subjects with both NEO data and quality-controlled MRI data. See sections 2.4 and

2.5 for further preprocessing details that explain how final sample sizes were derived. Demographic percentages in the text represent the full samples.

2.3 Data

2.3.1 Personality and OCPD Measures

Revised NEO Personality Inventory (NEO-PI-R)

The NEO-PI-R (Costa & McCrae, 1992), a 240-item self-report measure of the Five Factor Model of Personality, was administered in all 3 studies (**Figure 2.1**). It generates dimensional scores for five personality factors (i.e., extraversion, neuroticism, openness, conscientiousness, agreeableness) with six lower-order facets per factor (for a total of 30 facets). Each facet includes eight items scored from 0 to 4. This results in facet scores ranging from 0-32 and factor scores ranging from 0-192.

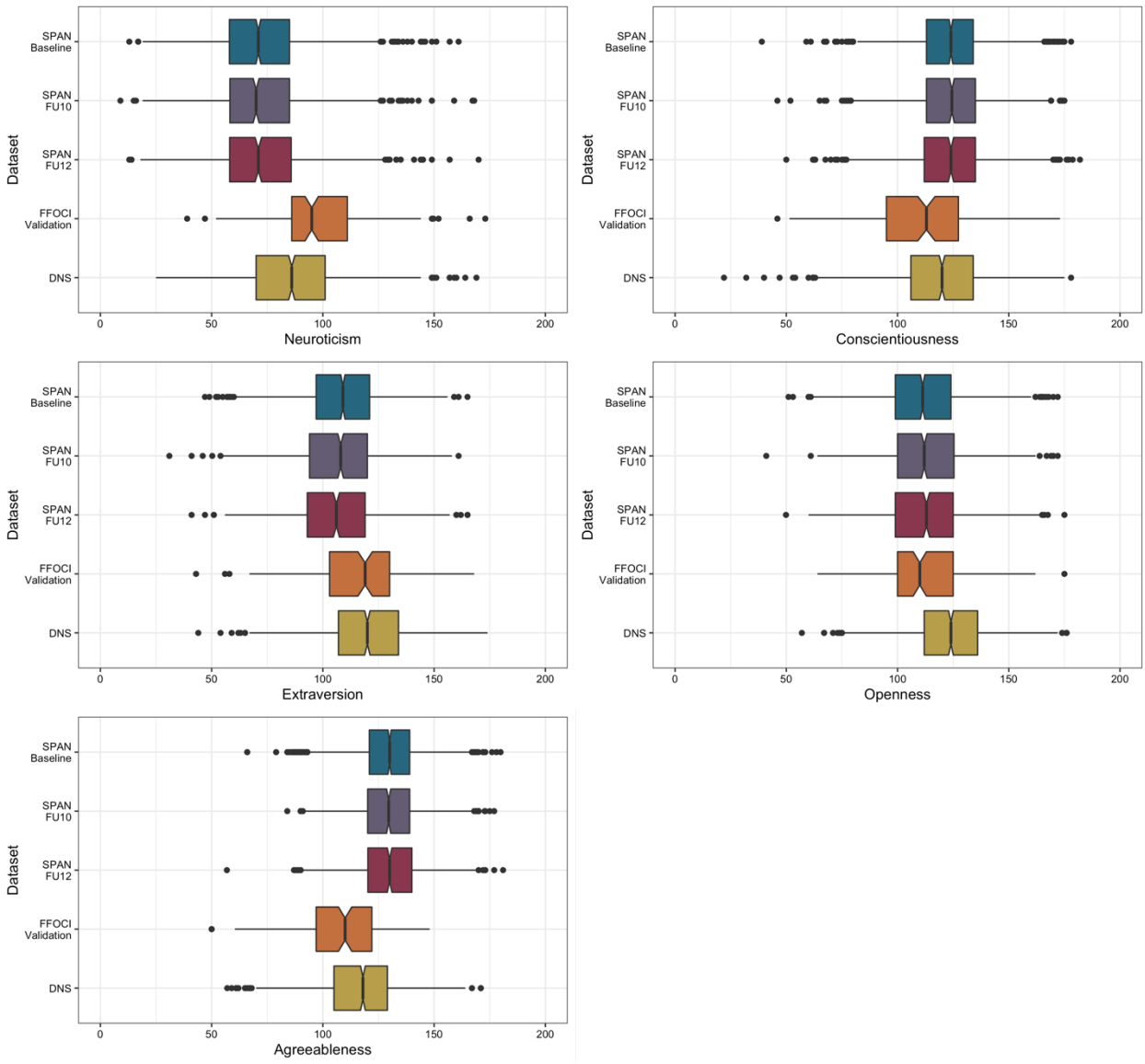


Figure 2.1 Personality Data from NEO-PI-R

Note: Data are from the final analytic samples and were preprocessed as described below.

Five-Factor Obsessive-Compulsive Inventory – Short Form (FFOCI-SF)

The FFOCI-SF (Griffin et al., 2018), a 48-item self-report measure of OCPD based on the Five Factor Model, was administered in the SPAN and FFOCI validation studies (**Figure 2.2**). It includes 12 scales that represent maladaptive variants of 12 facets of the FFM related to OCPD. Each scale is comprised of four items, scored from 1 to 5. In the present study, items were rescored on a 0-4 scale to be consistent with the NEO-PI-R scale (total score 0-192). The FFOCI-SF differs

slightly from the DSM-5 OCPD criteria in 1) including the historical notion of emotional inhibition and 2) excluding hoarding, which does not have a strong personality trait analog.

Structured Interview for DSM-IV Personality (SIDP-IV)

The SIDP-IV (Pfohl et al., 1997) is a clinician-administered semi-structured interview assessment of personality disorders. The eight diagnostic criteria for OCPD were scored based on interview responses on a scale from 0 (not present) to 3 (strongly present). At least four out of the eight criteria must be rated 2 ("present") or higher to meet diagnostic criteria for OCPD. In addition to a categorical OCPD variable, the sum of all OCPD responses (range: 0-24) was used as a dimensional score for OCPD (**Figure 2.2**). The SIDP-IV was only administered in the SPAN study, at baseline and follow-up 10. Interrater reliability for the baseline interviews was 0.67 (intraclass correlation) (Oltmanns et al., 2014). 46 participants (2.9%) met criteria for OCPD at baseline; at follow-up 10, 12 participants (0.9%) met criteria.

Multi-source Assessment of Personality Pathology (MAPP)

The MAPP (Oltmanns et al., 1998) is a 106-item questionnaire assessing normative personality (25 items) and personality disorders (81 items). Items are scored from 0 to 4. The OCPD scores include eight items that correspond to the eight OCPD diagnostic criteria in the DSM-IV. A dimensional score summed the scores of the eight OCPD items (range: 0-32; **Figure 2.2**), while categorical scores indicated the number of OCPD criteria that met diagnostic threshold (i.e., item scores of 3 or higher) and whether the individual met criteria for an OCPD diagnosis (i.e., four or more items had scores of 3 or higher). The MAPP was only administered in the SPAN study, at all three timepoints considered.

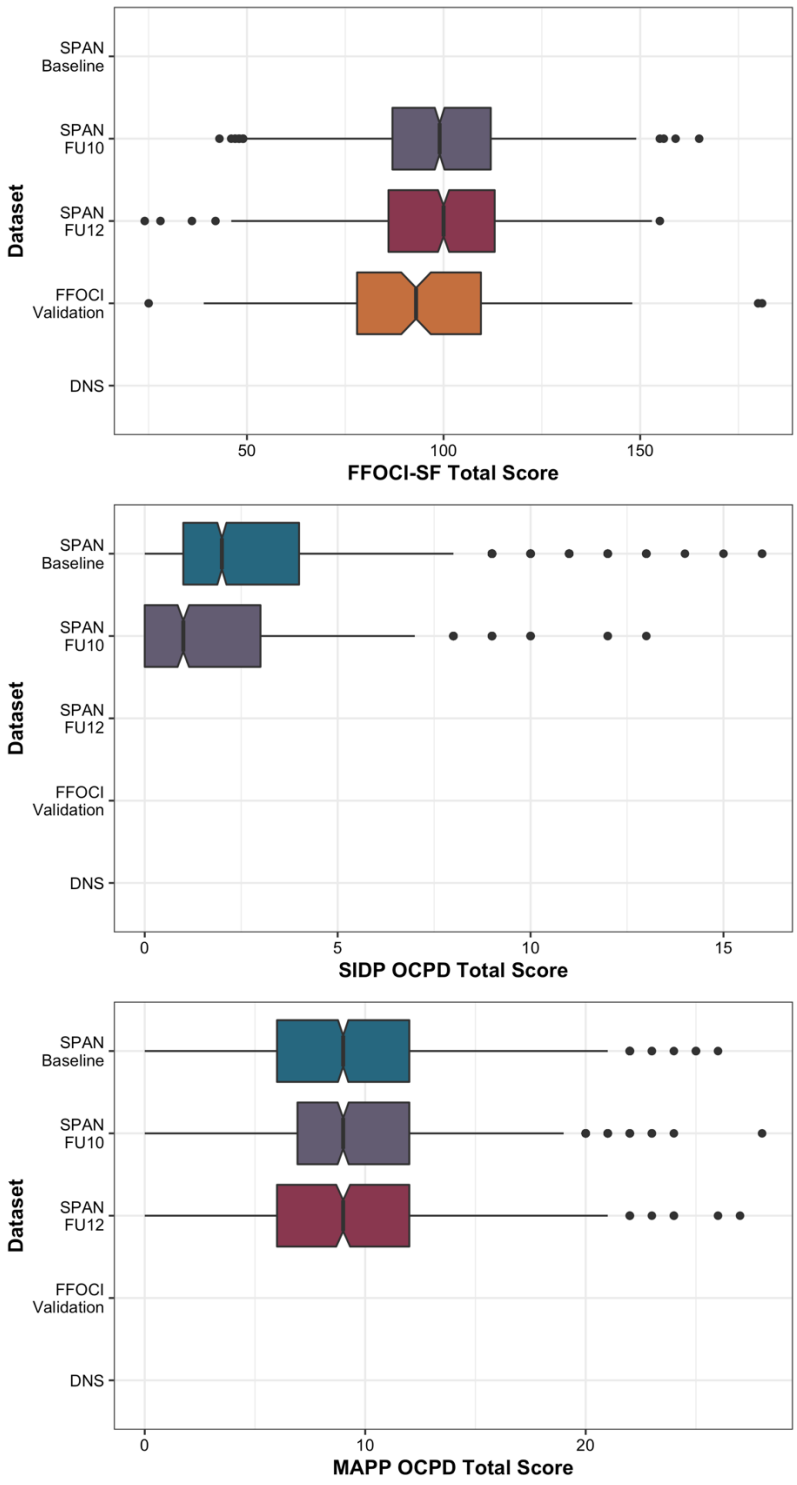


Figure 2.2 OCPD Data

Note: Data are from the final analytic samples and were preprocessed as described below, including scaling for missing items.

2.3.2 Correlates of OCPD

Depression
The Beck Depression Inventory II (BDI-II) (Beck et al., 1996), a 21-item dimensional measure of depressive symptoms (total score range: 0-63), was administered at all three SPAN timepoints.

Romantic Relationship Satisfaction
The Dyadic Adjustment Scale-4 (DAS-4) (Sabourin et al., 2005) was administered at all SPAN time points to assess participants' satisfaction with their romantic relationship, if applicable. Informant data from the romantic partner was also analyzed when available.

Social and Occupational Functioning
The Social Adjustment Scale-Self Report (SAS-SR) (Weissman, 1999) was administered at all SPAN time points. At baseline, the full questionnaire, except school functioning due to the sample's age range, was included to assess social (Section C) and occupational (Section A) functioning. Follow-up visits only included Section C on social functioning and leisure activities. The Scale of Unpleasant Relational Conduct Effects (SOURCE; (Lawton, 2014); total score range: 0-16) and the Quality of Relationship Inventory General Support subscale (QRI; (Pierce et al., 1991); total score range: 7-28) also provided data on SPAN participants' social functioning at follow-ups 10 and 12.

Education Attainment
SPAN participants reported their highest level of education obtained, which was analyzed as a categorical variable ranging from 1 (less than high school) to 9 (professional degree).

Life Satisfaction and Loneliness
The Satisfaction with Life (Diener et al., 1985) (5 items, total score range: 7-35) and UCLA Loneliness (Version 3) Scales (Russell, 1996) (20 items, total score range: 20-80) were administered at SPAN follow-ups 10 and 12.

Health Behaviors

The Health Behavior Checklist (Vickers et al., 1990) was administered to SPAN participants at follow-ups 10 and 12. The Good Health Practices scale (16 items; total score range: 16-80) was selected as a general measure of engaging in healthy behaviors.

Physical and Mental Health Help-Seeking

SPAN participants reported whether they had sought treatment or counseling for a psychiatric disorder or life problem (baseline = ever; FU10 and FU12 = in past six months). They also provided information on physical health treatment at follow-ups 10 and 12. For the present study, number of doctor's visits in past six months was used to assess this construct.

2.3.3 Neuroimaging Data

DNS participants were scanned on GE MR750 3 T MRI machines, using an eight-channel head coil. Of the 1,253 participants included in analyses (details below in 2.5.2), 1,035 participants were scanned on scanner one and 218 participants on scanner two. T1-weighted images were acquired using a 3D Axial FSPGR BRAVO sequence (*scan parameters:* TR = 8.148 ms; TE = 3.22 ms; 162 axial slices; flip angle = 12°; FOV = 240 mm; matrix = 256 x 256; slice thickness = 1 mm with no gap (voxel size 0.9375 x 0.9375 x 1 mm); scan time = 4 min and 13 s).

2.4 Aim 1 – OCPD Model

2.4.1 OCPD Predictive Model Training

The first aim of the present study was to train a machine learning model that could predict the score of an OCPD measure based on normative personality data. The goal was to be able to apply this trained model to a neuroimaging dataset without OCPD data in an effort to characterize brain structure correlates of OCPD traits. ML model training occurred in the SPAN dataset. Only participants with at least some data for all relevant measures (i.e., NEO, SIDP, MAPP, and FFOCI)

for a timepoint were considered (baseline=1,610; FU10=1,032; FU12=904) to facilitate model comparisons across different OCPD measures.

Preprocessing

Preprocessing was conducted to prepare the SPAN data for model training. First, the data were checked for near zero variance and/or highly correlated ($r>0.75$) predictor variables. None of the predictor variables (i.e., NEO items) met these criteria and required manipulation. Second, summary scores for OCPD measures were scaled to address missing data. Specifically, if enough items were completed, then the total scores were divided by the number of relevant items completed and multiplied by the expected number of items. If a participant was missing more than the accepted number of missing items, then the scaled score was marked as missing. Third, participants missing key data that were not addressed with the aforementioned scaling and that could not be reliably imputed were removed. Specifically, individuals with missing values for the scaled versions of the SIDP or MAPP OCPD item total scores or FFOCI-SF total were removed (baseline n=0; FU10 n=17; FU12 n=6) because too much data was missing for accurate values on the OCPD variables of interest. Individuals missing gender or age data (n=1) were also removed from the sample because these demographic variables were used to residualize the predictor and outcome variables for the OCPD models. Individuals missing data for more than 25 percent of NEO items (i.e., > 60 items; n=3 for baseline, 0 for FU10 and FU12) were removed because the ML models were trained on the individual NEO item data and imputation was unlikely to be accurate for participants missing that many NEO items. For the remaining participants, missing NEO item data was imputed with the individual's relevant facet mean. The final sample sizes were 1,606 (baseline), 1,015 (FU10), and 898 (FU12).

Fourth, data was split into training, validation, and testing datasets for each timepoint separately, using a 70/10/20 split. OCPD diagnosis, as defined by the SIDP (baseline and FU10) or MAPP (FU12), was used to balance group sizes when splitting data. Training datasets were used to test multiple ML algorithms on the candidate OCPD measures (i.e., SIDP OCPD items total score, MAPP OCPD items total score, and FFOCI-SF total score). Validation datasets were used to evaluate and compare the performances of the best algorithm for each OCPD measure at every timepoint. The testing dataset for the timepoint of the best-performing model in the validation data was used to estimate the model's out-of-sample performance.

Fifth, predictor (i.e., NEO items) and outcome (i.e., OCPD) variables were residualized with age (as a continuous variable) and gender (as a binary factor) to control for these demographics in the ML models. Residualizing was conducted separately for the training, validation, and testing datasets to prevent data leakage, when information is unintentionally shared between the training or testing data which leads to inflated model performance (Grzenda et al., 2021). All models were trained on both residualized and unresidualized data to assess the influence of controlling for these covariates and to determine if residualizing was necessary.

Finally, predictor variables were standardized (i.e., mean centered and divided by the standard deviation) within the model fitting and cross-validation process, instead of separately during pre-processing, to prevent data leakage.

Machine Learning Model Specification

Machine learning analyses were conducted with the caret package (version 6.0-91; (Kuhn, 2008)) in R (version 4.1.3). ML models were trained to predict OCPD scores (i.e., SIDP OCPD items total score, MAPP OCPD items total score, or FFOCI-SF total score) from normative personality data (i.e., all 240 NEO-PI-R items). Repeated 5x5 cross-validation (CV) was conducted in the

training data for hyperparameter tuning. Adaptive resampling, rather than grid or random search, was used for more efficient hyperparameter tuning, testing 100 hyperparameter combinations for the smallest root mean squared error (RMSE). Four supervised machine learning algorithms were tested, including elastic net regression ("glmnet" method in caret), gradient boosting machines ("gbm"), support vector regression with a linear kernel ("svmLinear"), and support vector regression with a radial kernel ("svmRadial"). Elastic net regression (Zou & Hastie, 2005) is a modified form of standard regression that uses regularization, a technique to decrease the variance of parameter estimates, which is often large in ordinary least squares regression when there are many predictor variables. Gradient boosting machines (J. Friedman, 2001) combine multiple "weak" models sequentially in an ensemble, attempting to minimize the ensemble's error as it goes, to "boost" the end model's predictive power (Natekin & Knoll, 2013). Support vector regression (Drucker et al., 1997) determines a model's best fit line within a specified error threshold. SVR may be conducted in a higher dimensional space than the original data and uses a "kernel" function to map data between the dimensions.

The "glmnet" and "gbm" algorithms include feature selection procedures during model training; no external feature selection was conducted prior to the SVR model fittings. These algorithms were applied to each OCPD measure in each timepoint, for both residualized and unresidualized data, resulting in 56 models trained in the training data (**Figure 2.3**). Parallel processing was implemented with the doParallel R package (version 1.0.17). Random number seeds and cross-validation indices, which specify the training data rows to use for each CV fold, were explicitly set with the trainControl function to ensure consistent and reproducible results from parallel processing. The caretList function was employed to run multiple models at once.

Because the OCPD measures were on different scales, the best performing ML algorithm for each OCPD measure and timepoint combination was identified by the ratio of the root mean squared error (RMSE) to outcome variable standard deviation. RMSE is the standard deviation of the residuals, indicating the absolute fit of the model to the data. It also provides the average model prediction error in the units of the outcome variable. Lower RMSE values indicate better fitting models. The RMSE value used was the average RMSE from cross-validation. Another common performance metric, mean absolute error (MAE), calculates the average absolute error between observed and predicted values. Although MAE was not used for model comparisons, MAE values are reported for model performance in the Results.

These best-performing trained models were then applied to the held-out validation data, generating predicted OCPD scores (**Figure 2.3**). Model performance was compared again, using the same metric as the training data, and the best overall model was selected. This final model was then applied to the held-out test data to generate predicted OCPD scores and obtain model performance metrics. This three-tiered process was implemented to identify the OCPD measure best predicted by the FFM and the best model while accounting for overfitting in training data. Caret's varImp function was run to examine feature importance.

2.4.2 External Validation of Trained OCPD Model

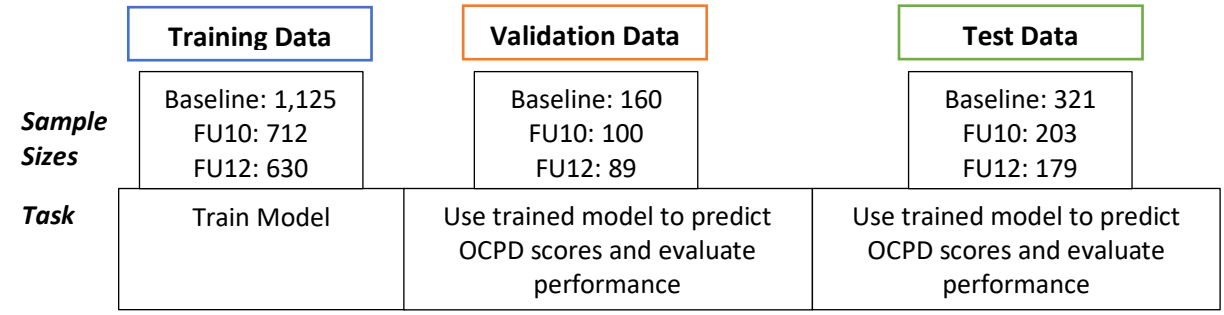
A second dataset with personality and OCPD data was used to test the external validity of the trained OCPD model and its generalizability to a college age sample. The data were preprocessed in the same way as the SPAN dataset used for model training (described above). In this dataset, 0 participants were missing > 25 % of NEO items and had to be removed; 5 participants were missing FFOCI-SF scaled totals and removed from the sample; and 23 participants were missing age or

gender data and also removed from the sample. The final sample for analyses included 175 participants.

Predicted FFOCI-SF scaled total scores were generated from the trained OCPD model. Model performance was assessed using RMSE, R^2, and the ratio of RMSE to the standard deviation of the observed scores.

2.4.3 Control Analyses: Correlates of OCPD

OCPD is often comorbid with depression (Grant et al., 2012) and associated with decreased marital satisfaction (South et al., 2020). Control analyses examined whether the observed and predicted OCPD scores in the SPAN dataset were correlated with these variables. They also investigated correlations between OCPD and social and occupational functioning, education attainment, life satisfaction, loneliness, health behaviors, and physical and mental health seeking behaviors.


Training Data
Validation Data
Test Data
Sample Sizes
Baseline: 1,125
FU10: 712
FU12: 630
Baseline: 160
FU10: 100
FU12: 89
Baseline: 321
FU10: 203
FU12: 179
Task
Train Model
Use trained model to predict OCPD scores and evaluate performance
Use trained model to predict OCPD scores and evaluate performance

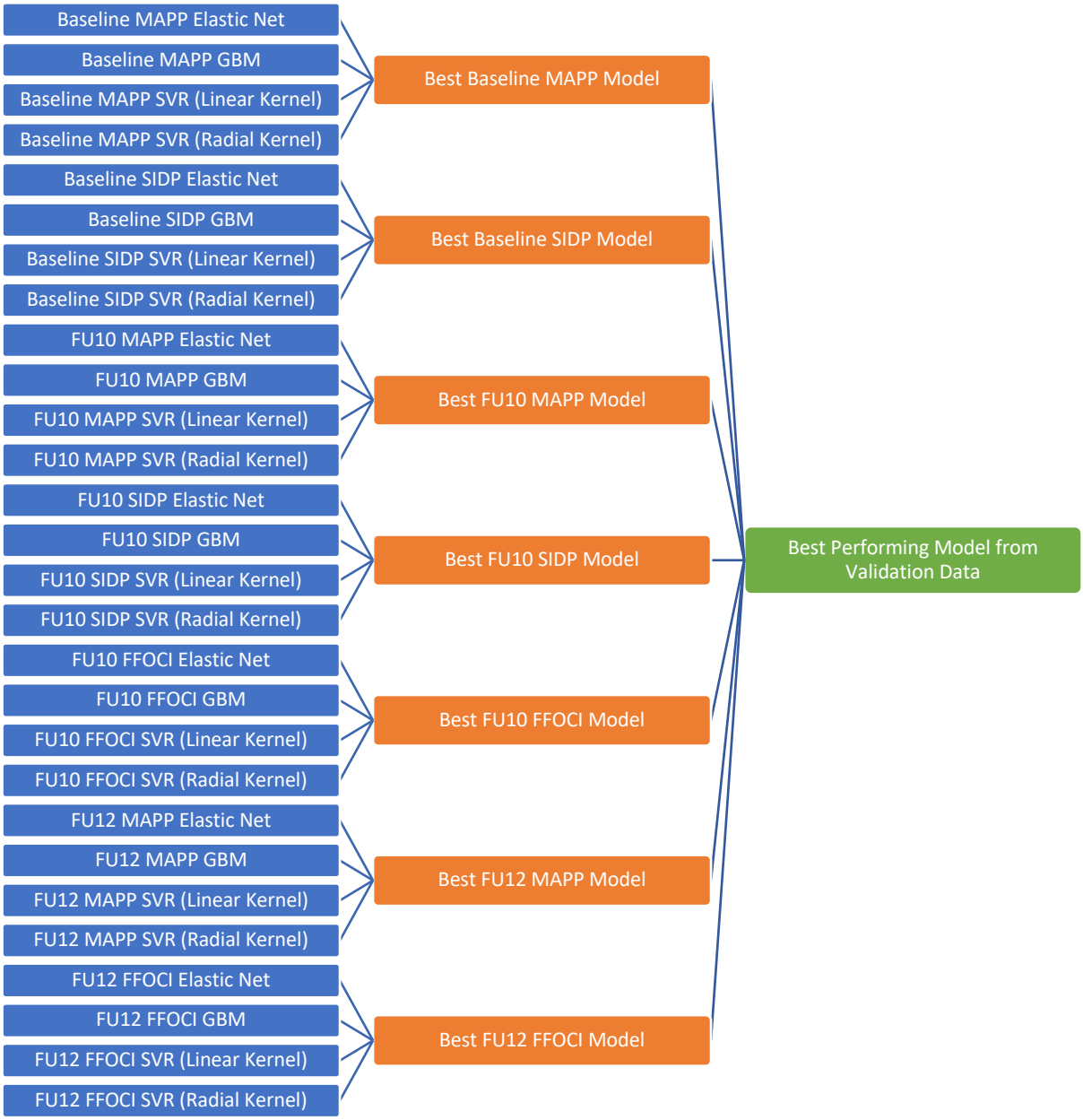

Baseline MAPP Elastic Net
Baseline MAPP GBM
Baseline MAPP SVR (Linear Kernel)
Baseline MAPP SVR (Radial Kernel)
Best Baseline MAPP Model
Baseline SIDP Elastic Net
Baseline SIDP GBM
Baseline SIDP SVR (Linear Kernel)
Baseline SIDP SVR (Radial Kernel)
Best Baseline SIDP Model
FU10 MAPP Elastic Net
FU10 MAPP GBM
FU10 MAPP SVR (Linear Kernel)
FU10 MAPP SVR (Radial Kernel)
Best FU10 MAPP Model
FU10 SIDP Elastic Net
FU10 SIDP GBM
FU10 SIDP SVR (Linear Kernel)
FU10 SIDP SVR (Radial Kernel)
Best FU10 SIDP Model
FU10 FFOCI Elastic Net
FU10 FFOCI GBM
FU10 FFOCI SVR (Linear Kernel)
FU10 FFOCI SVR (Radial Kernel)
Best FU10 FFOCI Model
FU12 MAPP Elastic Net
FU12 MAPP GBM
FU12 MAPP SVR (Linear Kernel)
FU12 MAPP SVR (Radial Kernel)
Best FU12 MAPP Model
FU12 FFOCI Elastic Net
FU12 FFOCI GBM
FU12 FFOCI SVR (Linear Kernel)
FU12 FFOCI SVR (Radial Kernel)
Best FU12 FFOCI Model
Best Performing Model from Validation Data

Figure 2.3. Machine Learning Model Selection Process

Each model in the training data (blue boxes) was run twice, once with unresidualized data and once with residualized data. The same algorithms performed best for each OCPD measure/timepoint combination, regardless of whether the data was residualized or not. Thus, the residualized and unresidualized versions of the best models from the training data were evaluated in the validation and test data.

2.5 Aim 2 – Neuroimaging Analyses

2.5.1 Generating Predicted OCPD Scores in Neuroimaging Dataset

The second aim of the present study was to characterize brain structure correlates of OCPD traits. We did not have access to a neuroimaging dataset with OCPD data, as personality disorder assessments are rarely conducted in large neuroimaging studies. Therefore, the trained OCPD ML model from Aim 1 was applied to the DNS neuroimaging dataset, which does not include OCPD measures, to generate predicted OCPD scores from normative personality data (i.e., the NEO-PI-R). Data were preprocessed in the same manner as the SPAN data used to train the model (described above). No participants were missing data for age or gender, or more than 25% of NEO items and had to be removed from the sample. As such, predicted OCPD scores were generated for the full dataset and the DNS dataset was ready for neural analyses.

2.5.2 Neuroimaging Data Processing

T1-weighted images were processed with the FreeSurfer image analysis suite (v5.3; http://surfer.nmr.mgh.harvard.edu/) to estimate volume, cortical thickness, and surface area (Dale et al., 1999; Fischl et al., 1999, 2002). Cortical regions were defined by the Desikan-Killiany-Tourville atlas (Klein & Tourville, 2012). Of the 1,321 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. One participant with acceptable neuroimaging data did not complete the NEO-PI-R and was thus missing a predicted FFOCI-SF score, so they were also excluded. The final sample size for gray matter analyses was 1,253.

2.5.3 Standard Regression Analyses

Neuroimaging analyses were first conducted with standard regression models. The volume, cortical thickness, or surface area of individual brain regions were standardized (i.e., centered and scaled) and then modeled as the independent variable of interest. Separate models were run for each brain region, hemisphere, and structural phenotype combination. Predicted FFOCI-SF total score (residualized and scaled for missing data, as described above) was modeled as the dependent variable. Covariates included age, gender (as factor), self-reported race (as factor), scanner (as factor), and global brain structure when analyzing a regional, rather than global, ROI (estimated total intracranial volume for volume and surface area models and average whole brain cortical thickness for cortical thickness models).

First, *a priori* analyses examined brain regions hypothesized to be related to OCPD (n=15 per hemisphere, 30 total): medial and lateral orbitofrontal cortex, dorsal anterior cingulate ("caudal anterior cingulate" in FreeSurfer), rostral anterior cingulate, medial and lateral prefrontal cortices (including the superior frontal and rostral and caudal middle frontal gyri (which contain the dlPFC) and inferior frontal (pars orbitalis, pars triangularis, pars opercularis) regions), as well as the insula, posterior cingulate cortex, precuneus, caudate, and ventral striatum (FreeSurfer's "Accumbens

Area"). Cortical thickness and surface area were examined for cortical regions while volume was examined for subcortical regions, resulting in 56 regression models. Multiple testing was corrected for with matrix spectral decomposition (MatSpD) (Nyholt, 2004) which accounts for the correlation between brain structure metrics to estimate the number of independent tests conducted. The volume of the subcortical ROIs and cortical thickness and surface area of the cortical ROIs were included together in the correlation matrix. The estimated number of independent tests (n=35) and subsequent alpha level (α=0.001464) was calculated using the method from (Li & Ji, 2005).

Next, exploratory analyses examined cortical thickness and surface area in all cortical ROIs in the Desikan-Killiany-Tourville atlas (n=31 regions per hemisphere, two hemispheres, and two phenotypes per ROI=124 regressions) and volume in all subcortical ROIs from FreeSurfer's automatic segmentation ("aseg") procedure *except* the ventricles, vessels, surface holes, hypointensities, cerebral spinal fluid, choroid plexus, optic chiasm, and brainstem (n=25 total; left and right hemisphere run separately when applicable). For global volume measures, cortical gray matter (total, left, and right), cerebral white matter (total, left, and right), subcortical gray matter, and total gray matter were included (n=8). Estimated total intracranial volume was not included because it was used as a covariate to control for global brain structure rather than a ROI of interest. These measurements were included in a separate correlation matrix to estimate the number of independent tests conducted (n=86.1) and subsequent alpha level (α=0.000595) for the 157 exploratory regression models.

2.5.4 Machine Learning Analyses

Next, neuroimaging analyses were conducted using machine learning techniques, which have been proposed to better capture the expected small effect sizes of psychopathology-related individual differences in morphometry (Dick et al., 2021; Götz et al., 2022; Walter et al., 2019b). Given

machine learning algorithms' ability to handle multiple predictor variables and capture complex interactions amongst them, all 157 measurements from the exploratory analyses described above were included as features in model training. The neuroimaging ML processing pipeline was largely consistent with the Aim 1 pipeline with a few exceptions. First, features (i.e., morphometric measures) were residualized for age and gender, like the OCPD models, but also for self-reported race and MRI scanner (which have known effects on brain morphometry (Tanga et al., 2010) and imaging output (Han et al., 2006), respectively) and global brain structure (estimated total intracranial volume for volume and surface area measurements and average whole brain cortical thickness for cortical thickness measurements). Second, data were randomly split into training (70%) and testing (30%) subsets, instead of training, validation, and testing datasets. Third, only participants with quality-controlled FreeSurfer data were included, so there was no missing feature data (i.e., morphometric measures) that required imputation. Fourth, a neural net algorithm was also trained (method "nnet" in caret from "nnet" R package), in addition to the elastic net, gradient boosting machines, and linear and radial support vector regression algorithms employed in Aim 1.

Chapter 3: Results

3.1 Aim 1 – OCPD ML Model Training

3.1.1 Final Model Performance

Broadly, the FFM items in the NEO-PI-R were better able to predict scores on the FFOCI-SF than the MAPP and SIDP across all time points (**Table 3.1**). In the validation dataset, the ratio of RMSE to standard deviation was ~0.6 for FFOCI, ~0.85 for MAPP, and ~0.87 for SIPD. An elastic net regression algorithm and FFOCI-SF data at follow-up 12 had the lowest ratio of RMSE (i.e., error) to standard deviation in the validation data (0.523), and thus produced the best model fit (**Table 3.1;** *testing data:* RMSE=12.07, R^2=0.55, MAE=9.6; observed SD=18.09; n=179; **Figure 3.1**). Hyperparameter tuning, conducted using adaptive resampling in caret, identified the optimal values of alpha (0.0921), the mixing percentage, and lambda (5.078), the regularization parameter. In elastic net regression, the mixing percentage determines the relative weights, or "mixture", of lasso and ridge penalties (here 9.21% lasso and 90.79% ridge); these penalties are based on the size of the coefficients and aim to decrease the risk of overfitting. The regularization parameter determines the size of that penalty, with higher values indicating fewer variables will be kept in the model. The regularization feature selection process retained 114 of the 240 NEO items in the model (**Figure 3.2**). Performance for the unresidualized model was very similar (RMSE=12.41, R^2=0.56, MAE=9.60, observed SD=18.32, RMSE/SD=0.68, 113 features; **Table 3.1**).

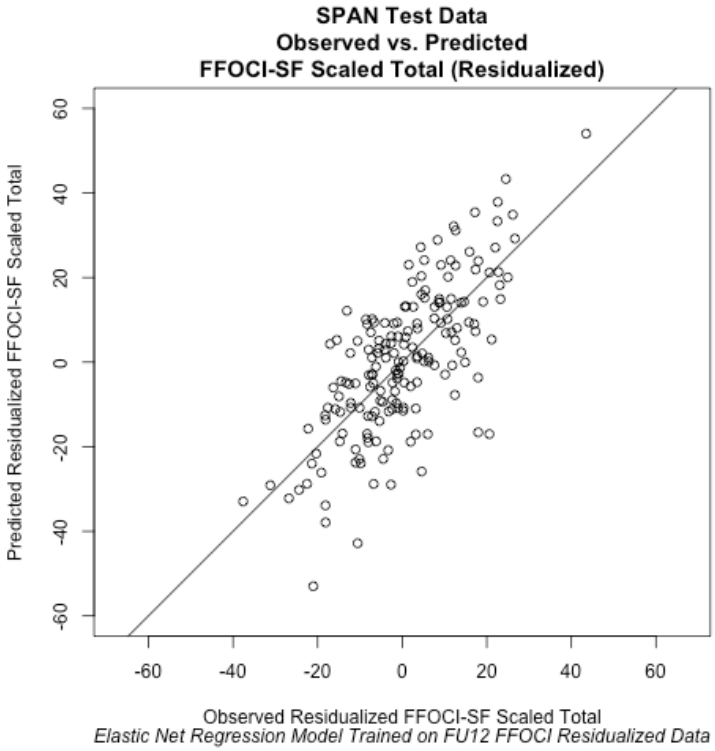


Figure 3.1. Actual versus predicted FFOCI-SF residualized values in the SPAN testing dataset for the final OCPD model

Table 3.1. OCPD Model Training Results

Time Point	Outcome Measure	Algorithm	Unresidualized - *Training* Data					Residualized - *Training* Data					Unresidualized - *Validation* Data					Residualized - *Validation* Data				
			RMSE	R^2	MAE	SD	RMSE/SD	RMSE	R^2	MAE	SD	RMSE/SD	RMSE	R^2	MAE	SD	RMSE/SD	RMSE	R^2	MAE	SD	RMSE/SD
Baseline	MAPPOC Scaled Total	glmnet	3.76	0.31	3.01	4.52	0.832	3.75	0.29	3.00	4.46	0.842	---	---	---	---	---	---	---	---	---	---
Baseline	MAPPOC Scaled Total	gbm	3.84	0.29	3.06	4.52	0.850	3.82	0.27	3.03	4.46	0.856	---	---	---	---	---	---	---	---	---	---
Baseline	MAPPOC Scaled Total	svmLinear	4.33	0.22	3.43	4.52	0.959	4.32	0.20	3.42	4.46	0.968	---	---	---	---	---	---	---	---	---	---
Baseline	**MAPPOC Scaled Total**	**svmRadial**	**3.75**	**0.32**	**2.98**	**4.52**	**0.829**	**3.73**	**0.30**	**2.97**	**4.46**	**0.837**	**3.92**	**0.28**	**3.02**	**4.60**	**0.852**	**3.90**	**0.27**	**3.01**	**4.59**	**0.851**
Baseline	**SIDPOC Scaled Total**	**glmnet**	**2.54**	**0.15**	**1.91**	**2.75**	**0.924**	**2.53**	**0.14**	**1.91**	**2.72**	**0.929**	**2.23**	**0.19**	**1.71**	**2.48**	**0.899**	**2.21**	**0.21**	**1.68**	**2.47**	**0.892**
Baseline	SIDPOC Scaled Total	gbm	2.56	0.14	1.94	2.75	0.931	2.56	0.12	1.94	2.72	0.941	---	---	---	---	---	---	---	---	---	---
Baseline	SIDPOC Scaled Total	svmLinear	2.83	0.09	2.10	2.75	1.032	2.83	0.08	2.09	2.72	1.038	---	---	---	---	---	---	---	---	---	---
Baseline	SIDPOC Scaled Total	svmRadial	2.54	0.14	1.94	2.75	0.927	2.55	0.13	1.94	2.72	0.935	---	---	---	---	---	---	---	---	---	---
FU10	**MAPPOC Scaled Total**	**glmnet**	**3.62**	**0.32**	**2.86**	**4.40**	**0.823**	**3.61**	**0.31**	**2.85**	**4.36**	**0.829**	**3.82**	**0.29**	**2.97**	**4.52**	**0.845**	**3.80**	**0.24**	**2.94**	**4.35**	**0.872**
FU10	MAPPOC Scaled Total	gbm	3.64	0.32	2.89	4.40	0.826	3.70	0.28	2.94	4.36	0.850	---	---	---	---	---	---	---	---	---	---
FU10	MAPPOC Scaled Total	svmLinear	4.39	0.20	3.50	4.40	0.998	4.38	0.19	3.49	4.36	1.006	---	---	---	---	---	---	---	---	---	---
FU10	MAPPOC Scaled Total	svmRadial	3.62	0.33	2.86	4.40	0.823	3.62	0.31	2.85	4.36	0.832	---	---	---	---	---	---	---	---	---	---
FU10	**SIDPOC Scaled Total**	**glmnet**	**1.87**	**0.19**	**1.42**	**2.08**	**0.899**	**1.86**	**0.19**	**1.42**	**2.07**	**0.899**	**1.69**	**0.26**	**1.29**	**1.96**	**0.862**	**1.67**	**0.27**	**1.25**	**1.96**	**0.851**
FU10	SIDPOC Scaled Total	gbm	1.88	0.19	1.43	2.08	0.904	1.91	0.16	1.46	2.07	0.922	---	---	---	---	---	---	---	---	---	---
FU10	SIDPOC Scaled Total	svmLinear	2.26	0.09	1.74	2.08	1.090	2.26	0.09	1.73	2.07	1.088	---	---	---	---	---	---	---	---	---	---
FU10	SIDPOC Scaled Total	svmRadial	1.90	0.17	1.44	2.08	0.913	1.89	0.17	1.44	2.07	0.913	---	---	---	---	---	---	---	---	---	---
FU10	**FFOCI Scaled Total**	**glmnet**	**12.37**	**0.57**	**9.77**	**18.80**	**0.658**	**12.33**	**0.57**	**9.77**	**18.70**	**0.659**	**13.65**	**0.61**	**10.39**	**20.75**	**0.658**	**13.05**	**0.61**	**10.16**	**20.52**	**0.636**
FU10	FFOCI Scaled Total	gbm	12.77	0.54	10.21	18.80	0.679	13.13	0.51	10.52	18.70	0.702	---	---	---	---	---	---	---	---	---	---
FU10	FFOCI Scaled Total	svmLinear	15.69	0.41	12.59	18.80	0.835	15.66	0.40	12.56	18.70	0.837	---	---	---	---	---	---	---	---	---	---
FU10	FFOCI Scaled Total	svmRadial	12.56	0.56	10.01	18.80	0.668	12.53	0.55	9.98	18.70	0.670	---	---	---	---	---	---	---	---	---	---
FU12	MAPPOC Scaled Total	glmnet	3.86	0.30	3.02	4.60	0.839	3.86	0.28	3.02	4.55	0.848	---	---	---	---	---	---	---	---	---	---
FU12	MAPPOC Scaled Total	gbm	4.00	0.25	3.16	4.60	0.870	3.96	0.25	3.10	4.55	0.869	---	---	---	---	---	---	---	---	---	---
FU12	MAPPOC Scaled Total	svmLinear	4.77	0.18	3.77	4.60	1.037	4.77	0.17	3.77	4.55	1.048	---	---	---	---	---	---	---	---	---	---
FU12	**MAPPOC Scaled Total**	**svmRadial**	**3.85**	**0.30**	**3.01**	**4.60**	**0.837**	**3.85**	**0.29**	**3.02**	**4.55**	**0.845**	**3.59**	**0.31**	**2.86**	**4.33**	**0.830**	**3.58**	**0.28**	**2.84**	**4.23**	**0.846**
FU12	**FFOCI Scaled Total**	**glmnet**	**12.02**	**0.62**	**9.55**	**19.53**	**0.615**	**12.02**	**0.62**	**9.55**	**19.48**	**0.617**	**10.32**	**0.76**	**8.20**	**19.93**	**0.518**	**10.18**	**0.75**	**8.18**	**19.45**	**0.523**
FU12	FFOCI Scaled Total	gbm	12.61	0.58	9.98	19.53	0.646	12.77	0.58	10.14	19.48	0.656	---	---	---	---	---	---	---	---	---	---
FU12	FFOCI Scaled Total	svmLinear	14.88	0.49	11.66	19.53	0.762	15.01	0.49	11.69	19.48	0.770	---	---	---	---	---	---	---	---	---	---
FU12	FFOCI Scaled Total	svmRadial	12.13	0.62	9.54	19.53	0.621	12.14	0.61	9.53	19.48	0.623	---	---	---	---	---	---	---	---	---	---

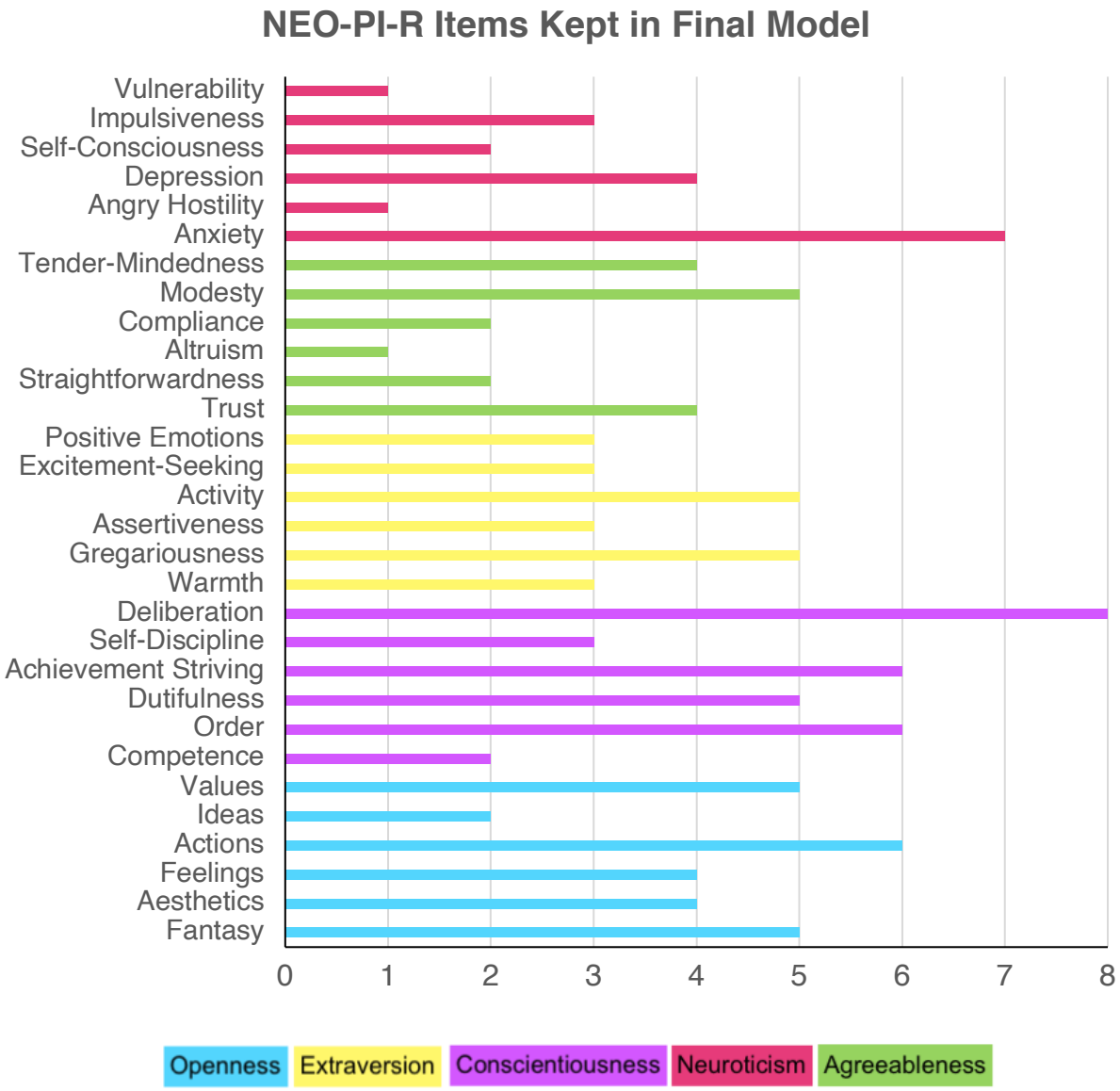


Figure 3.2 NEO-PI-R items kept in model by facet

The residualized trained model performed similarly well in the external validation dataset (RMSE=12.06, R^2=0.75, MAE=8.92; observed SD=23.56, RMSE/SD=0.51; n=175; **Figure 3.3**), suggesting evidence of model generalizability. Results were nearly identical when the two age outliers were excluded, so they were retained in the sample. The unresidualized model performed worse than the residualized model (RMSE=13.70, R^2=0.75, MAE=10.41, observed SD=23.59, RMSE/SD=0.58; n=175). This suggested that residualizing the data to control for gender and age

was important to model performance, as the validation dataset was much younger than the dataset used to train the model (SPAN) but the same age range as the dataset the model was designed to be applied to (DNS). Thus, the residualized version of the trained model was applied to the DNS dataset in Aim 2 (described below).

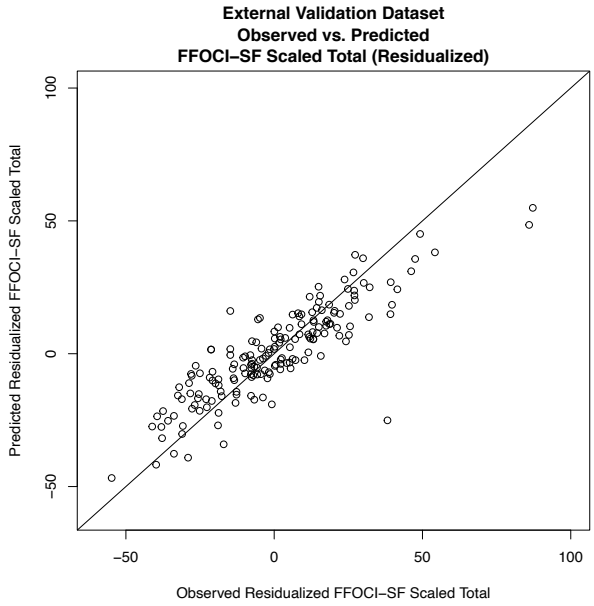


Figure 3.3. Actual versus predicted FFOCI-SF residualized values in the FFOCI Study external validation dataset using the final trained OCPD model

3.1.2 Feature Importance

Feature importance information, also known as variable importance, provides an indication of the relative importance of included predictors. Multivariate predictive models are dependent upon all of the input features and therefore individual features should not be strongly interpreted in isolation (Nielsen et al., 2020). However, this information can still provide insights into which predictor variables (in the current study, NEO items) were influential in model performance. For the final model (an elastic net regression model), feature importance was determined by the regression coefficients. The distribution of feature importance values steadily decreased from the highest

(1.69) to the lowest (0.00009); a cluster of most important items was not readily apparent. Conscientiousness had the highest average ranking of individual items' feature importance values (50.1 out of 114, 63% of items kept in model), followed by openness (54.3 out of 114, 54% of items kept in model), extraversion (58.7 out of 114, 46% of items kept in model), neuroticism (62.1 out of 114, 38% of items kept in model), and agreeableness (68.5 out of 114, 38% kept in model). The top 20 most important NEO items (**Figure 3.4**) were 35% conscientiousness, 30% openness, 15% extraversion, 10% neuroticism, and 10% agreeableness. These items came from the following facets: conscientiousness - achievement striving (3 items), deliberation (1 item), dutifulness (1 item), order (2 items); openness - actions (2 items), fantasy (1 item), and values (3 items); extraversion - gregariousness (2 items), and warmth (1 item); agreeableness - tender-mindedness (1 item) and trust (1 item); neuroticism - anxiety (1 item) and self-consciousness (1 item).

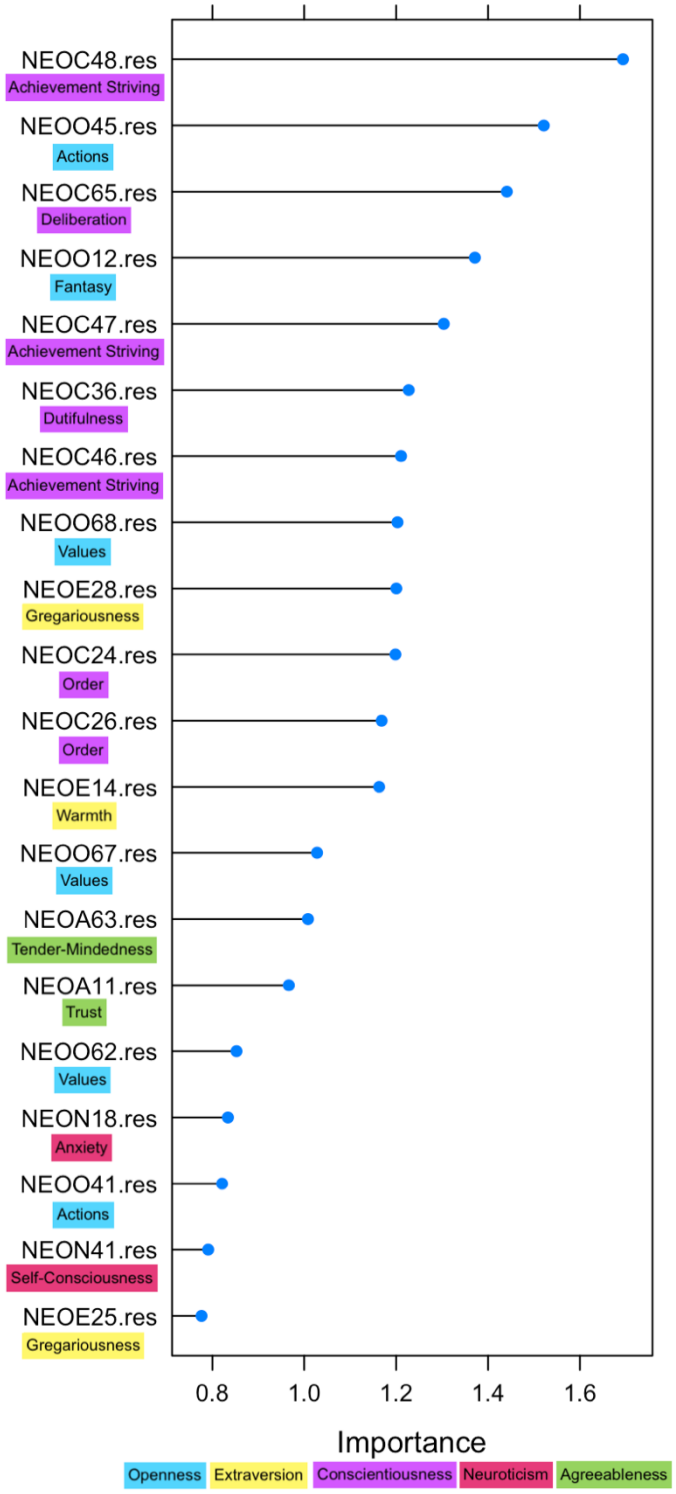


Figure 3.4. Feature Importance for Final Trained OCPD Model

Note: Values represent regression coefficients from the residualized elastic net regression model for FU12 FFOCI data.

3.1.3 Control Analyses: Correlates of OCPD

OCPD was correlated with increased depression severity and loneliness but also marginally better occupational functioning. OCPD was also associated with decreased romantic relationship and life satisfaction and engagement in healthy behaviors. The direction of associations was inconsistent across the OCPD measures (i.e., MAPP, SIDP, and FFOCI) for education attainment, social functioning, and physical and mental health care received. Notably, the strengths of the associations were very similar for the predicted FFOCI-SF scores from the final ML model and the corresponding observed scores (**Table 3.2**).

Table 3.2 Correlates of OCPD

	Measure	Baseline		FU10			FU12		
		MAPP	SIDP	MAPP	SIDP	FFOCI	MAPP	FFOCI	*Predicted* FFOCI
Occupational Functioning	SAS Factor A	**0.070**	0.058						
Social Functioning	SAS Factor C	**0.23*** **	**0.086** **	**0.17*** **	**0.20*** **	**0.26*** **	**0.21*** **	**0.23*** **	**0.23*** **
	SOURCE - *Participant*			**0.25*** **	**0.23*** **	**0.15*** **	**0.20*** **	**0.15*** **	**0.18*** **
	SOURCE - *Informant*			0.062	**0.10***	0.039	**0.16*** **	0.019	0.060
	QRI			**-0.15*** **	**-0.17*** **	**-0.19*** **	**-0.17*** **	**-0.18*** **	**-0.17*** **
Education Level	Categories	0.014	0.044	**0.072**	-0.047	**-0.12*** **	0.031	**-0.097***	**-0.16*** **
Depression	BDI-II	**0.27*** **	**0.13*** **	**0.26*** **	**0.26*** **	**0.20*** **	**0.27*** **	**0.16*** **	**0.20*** **
Relationship Satisfaction	DAS - *Participant*	**-0.087***	-0.059	**-0.14** **	**-0.15** **	**-0.19*** **	**-0.11**	**-0.086**	**-0.089**
	DAS - *Partner*	-0.073	-0.067	-0.063	-0.088	-0.048	**-0.12**	0.001	-0.087
Life Satisfaction	SWLS			**-0.15*** **	**-0.21*** **	**-0.18*** **	**-0.13*** **	**-0.078**	**-0.098***
Loneliness	UCLA			**0.29*** **	**0.29*** **	**0.24*** **	**0.30*** **	**0.19*** **	**0.23*** **
Health Behaviors	HBC GHP16			-0.049	-0.057	**-0.065**	**-0.073**	**-0.10***	**-0.098***
Medical Care	# Recent Doctor Appts			**0.072**	**0.12** **	0.010	0.014	-0.065	-0.048
Help-Seeking	Received MH Tx (Yes/No)	-0.036	-0.025	-0.024	0.12	-0.029	0.10	0.029	0.069

Notes: Help-seeking correlations are point-biserial correlations; all other values are Pearson correlation coefficients. The color gradient indicates strongest negative to strongest positive correlations. Bold=p<0.05, *=p<0.01, **=p<0.001, ***=p<0.0001, SAS=Social Adjustment Scale, SOURCE=Scale of Unpleasant Relational Conduct Effects, QRI=Quality of Relationship Inventory-General Support Subscale, BDI-II=Beck Depression Inventory–II, DAS=Dyadic Adjustment Scale, SWLS=Satisfaction with Life Scale, UCLA=UCLA Loneliness Scale (Version 3), HBC=Health Behavior Checklist, GHP16=Good Health Practices-16, MH=Mental Health, Tx=Treatment, MAPP=Multisource Assessment of Personality Pathology, SIDP=Structured Interview for DSM-IV Personality (SIDP-IV), FFOCI=Five-Factor Obsessive Compulsive Inventory-Short Form, Predicted FFOCI=FFOCI-SF

residualized predicted scores generated from the final OCPD ML model in Aim 1. All measures except occupational functioning, education level, number of recent doctor appointments and mental health treatment were scaled to address missing data.

3.2 Aim 2 – Neuroimaging Analyses

3.2.1 Standard Regressions

***A priori* Regions of Interest Analyses**

Right superior frontal cortical thickness was significantly associated with the predicted OCPD score (i.e., FFOCI-SF total score scaled for missing data and residualized) wherein thicker cortex was associated with higher OCPD scores (b= 2.21 [95% Confidence Interval (CI): 0.85-3.57], t(1,242)=3.19, p=0.001435). Regions with nominally significant results that did not survive correction for multiple comparisons included the left insula (b=-1.33, p=0.0099; cortical thickness) and pars triangularis (b=-1.15, p=0.047; cortical thickness) and right precuneus (b=-1.67, p=0.014; surface area) and pars orbitalis (b=-1.1, p=0.048; surface area).

Exploratory Analyses

No additional brain regions exhibited a significant association between brain structure and predicted OCPD score. Nominally significant results were found for surface area in the left fusiform gyrus, middle temporal gyrus, cuneus, and entorhinal cortex.

3.2.2 Machine Learning

The elastic net regression model performed best in the MRI training data (n=880; **Table 3.3**). Hyperparameter tuning, conducted using adaptive resampling in caret, identified the optimal values of alpha (0.92), the mixing percentage, and lambda (1.15), the regularization parameter. The regularization feature selection process retained 7 of the 157 morphometry measures in the model (**Figure 3.5**). Performance metrics from the testing data (n=373) were: RMSE=16.23, R^2=0.002, MAE=12.65; outcome SD=16.16; RMSE/SD=1.005.

Table 3.3. MRI training data model performance

Algorithm	RMSE	R^2	MAE	SD	RMSE/SD
glmnet	**16.067**	**0.00801**	**12.696**	**16.11**	**0.997**
gbm	16.237	0.00557	12.802	16.11	1.008
svmLinear	18.147	0.00418	14.294	16.11	1.126
svmRadial	16.107	0.00428	12.708	16.11	1.000
nnet	16.095	NA	12.698	16.11	0.999

Note: Both features and outcome variables were residualized. SD = outcome variable standard deviation.

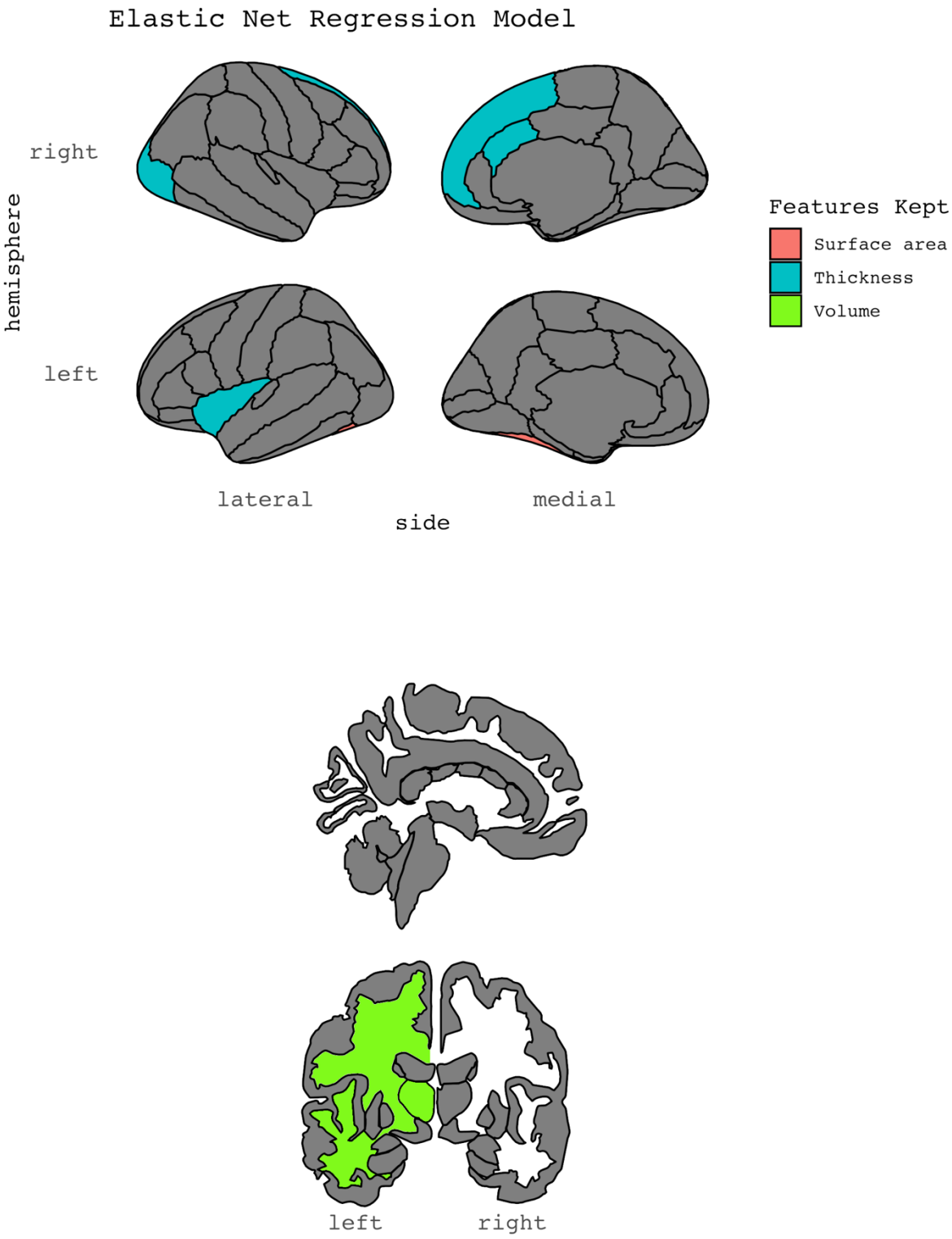


Figure 3.5. Morphometry measures retained in the machine learning model

Feature Importance

The seven morphometric measurements retained in the model included (from largest to smallest importance; **Figure 3.6**): left insula thickness, left thalamus volume, left fusiform gyrus surface area, right lateral occipital cortex thickness, left hemisphere cerebral white matter volume, right superior frontal gyrus thickness, and right caudal anterior cingulate thickness. However, given the high error rate of the model, interpretation of the feature importance data is severely limited.

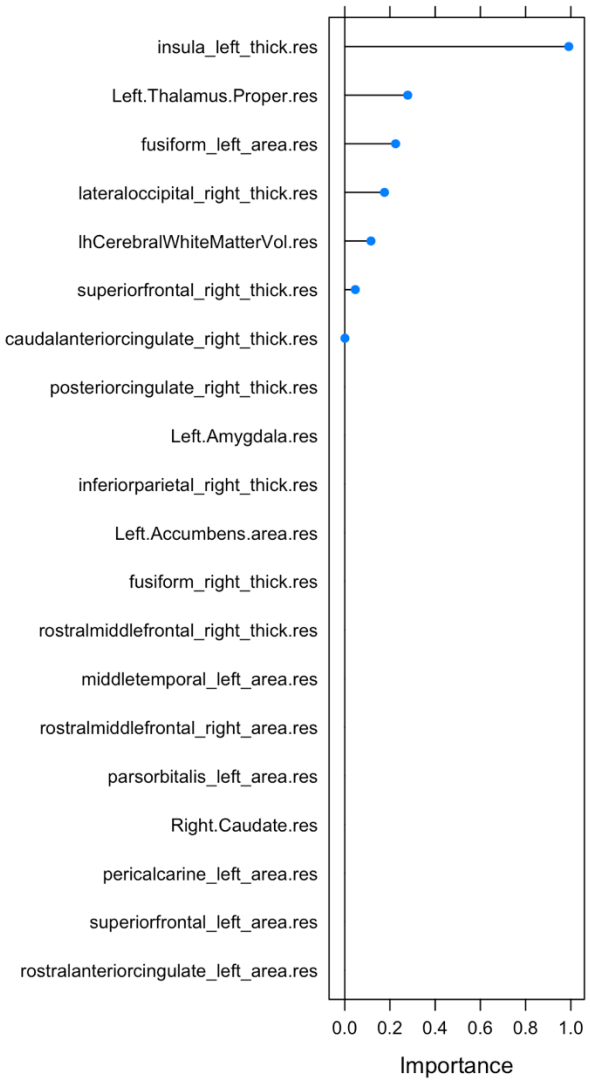


Figure 3.6. Feature Importance Values for the OCPD Neuroimaging Model

Post-Hoc Tests

Several follow-up tests were conducted in response to the poor model performance. First, models were trained using unresidualized outcome and/or features data, with and without covariates, to assess whether the residualization process was degrading performance. However, the performance metrics of these models were very similar to the residualized models (**Table 3.4)**, suggesting that residualizing was not causing the poor performance. Second, models were trained using only the a priori ROIs included in the a priori regression analyses to assess whether including all ROIs was obscuring signal from important brain regions. Again, model performance was similar to the original analyses (**Table 3.5**). Third, models were trained to predict average cortical thickness from individual ROIs' thickness to test whether the ML algorithms were capable of capturing expected strong relationships. The residualized models continued to perform poorly, with the RMSE equal to the standard deviation of average thickness (**Table 3.6**). The unresidualized models performed well, with an RMSE one-third of the outcome standard deviation. This suggested that errors in data or model processing were not causing the poor performance, given that average thickness could be accurately predicted as expected.

Supplemental Post-Hoc Tests

Additional post-hoc tests were conducted to assess whether brain structure accurately predicted neuroticism and conscientiousness, two of the personality domains most strongly associated with OCPD. The same machine learning algorithms as the main analyses (i.e., elastic net, gradient boosting machines, support vector machine with linear kernel, support vector machine with radial kernel) were run. Here, the outcome variable of interest was the residualized neuroticism or conscientiousness domain score from the NEO-PI-R. Similar to the main analyses, the neuroimaging features did not reliably predict personality (training data: RMSE/outcome SD$\cong$ 1, $R^2 \cong 0.006$; **Table 3.7**).

Table 3.4 MRI Model Performance for Unresidualized Data

	Residualized?		Training Data					Testing Data				
Algorithm	**Brain data?**	**FFOCI -SF?**	**RMSE**	**R^2**	**MAE**	**Outcome SD**	**RMSE/ SD**	**RMSE**	**R^2**	**MAE**	**Outcome SD**	**RMSE/ SD**
No Covariates												
glmnet	No	Yes	16.064	0.010	12.684	16.11	0.997	16.167	0.001	12.610	16.16	1.001
gbm	No	Yes	16.220	0.006	12.802	16.11	1.007	16.229	0.001	12.619	16.16	1.004
svmLinear	No	Yes	17.879	0.005	14.093	16.11	1.110	18.109	0.000	14.202	16.16	1.121
svmRadial	No	Yes	16.088	0.005	12.688	16.11	0.999	16.148	0.001	12.528	16.16	0.999
nnet	No	Yes	16.085	0.009	12.683	16.11	0.998	16.129	0.001	12.532	16.16	0.998
glmnet	No	No	16.022	0.011	12.636	16.07	0.997	16.140	0.001	12.554	16.14	1.000
gbm	No	No	16.170	0.007	12.764	16.07	1.006	16.189	0.003	12.558	16.14	1.003
svmLinear	No	No	17.787	0.005	14.005	16.07	1.107	18.035	0.000	14.090	16.14	1.117
svmRadial	No	No	16.038	0.006	12.633	16.07	0.998	16.130	0.001	12.459	16.14	0.999
nnet	No	No	98.963	NA	97.652	16.07	6.157	99.651	NA	98.339	16.14	6.175
Covariates Included as Features												
glmnet	No	Yes	16.006	0.018	12.625	16.11	0.993	16.172	0.001	12.594	16.16	1.001
gbm	No	Yes	16.196	0.008	12.790	16.11	1.005	16.218	0.002	12.647	16.16	1.004
svmLinear	No	Yes	17.601	0.005	13.890	16.11	1.092	18.099	0.000	14.217	16.16	1.120
svmRadial	No	Yes	16.043	0.012	12.627	16.11	0.996	16.257	0.001	12.589	16.16	1.006
nnet	No	Yes	16.085	0.012	12.685	16.11	0.998	16.145	0.002	12.574	16.16	0.999
glmnet	No	No	15.968	0.019	12.574	16.07	0.993	16.151	0.0018	12.537	16.14	1.001
gbm	No	No	16.143	0.008	12.734	16.07	1.004	16.186	0.0024	12.553	16.14	1.003
svmLinear	No	No	17.556	0.005	13.849	16.07	1.092	18.068	0.0004	14.180	16.14	1.120
svmRadial	No	No	15.998	0.013	12.581	16.07	0.995	16.239	0.0008	12.533	16.14	1.006
nnet	No	No	98.963	NA	97.652	16.07	6.157	99.651	NA	98.339	16.14	6.175

Note: Models included all morphometry measures (i.e., exploratory). FFOCI-SF = Five-Factor Obsessive Compulsive Inventory – Short Form. RMSE = root mean square error. MAE = mean absolute error. SD = standard deviation

Table 3.5 MRI Model Performance: Exploratory vs. a priori ROIs

	Training Data					Testing Data				
Algorithm	**RMSE**	**R^2**	**MAE**	**Outcome SD**	**RMSE/ SD**	**RMSE**	**R^2**	**MAE**	**Outcome SD**	**RMSE/ SD**
Exploratory										
glmnet	16.067	0.00801	12.696	16.11	0.997	16.233	0.00193	12.653	16.157	1.005
gbm	16.237	0.00557	12.802	16.11	1.008	16.318	0.00003	12.666	16.157	1.01
svmLinear	18.147	0.00418	14.294	16.11	1.126	18.2	0.00002	14.155	16.157	1.126
svmRadial	16.107	0.00428	12.708	16.11	1.000	16.159	0	12.56	16.157	1
nnet	16.095	NA	12.698	16.11	0.999	16.145	NA	12.574	16.157	0.999
A priori										
glmnet	16.068	0.00987	12.690	16.11	0.997	16.177	0.00133	12.611	16.157	1.001
gbm	16.270	0.00380	12.801	16.11	1.010	16.269	0.00038	12.809	16.157	1.007
svmLinear	16.863	0.00460	13.312	16.11	1.047	16.656	0.00178	13.034	16.157	1.031
svmRadial	16.097	0.00613	12.706	16.11	0.999	16.141	0.00114	12.565	16.157	0.999
nnet	16.092	0.00715	12.693	16.11	0.999	16.166	0.00214	12.606	16.157	1.001

Note: Both MRI features and the FFOCI-SF outcome variable were residualized. RMSE = root mean square error. MAE = mean absolute error. SD = standard deviation.

Table 3.6 Model performance for predicting average cortical thickness

	Training Data					Testing Data				
Algorithm	**RMSE**	**R^2**	**MAE**	**Outcome SD**	**RMSE/ SD**	**RMSE**	**R^2**	**MAE**	**Outcome SD**	**RMSE/ SD**
Residualized										
glmnet	0.0747	NA	0.0595	0.075	1	0.0772	NA	0.0616	0.0773	0.999
gbm	0.0757	0.005	0.0603	0.075	1.01	0.0771	0.0060	0.0618	0.0773	0.997
svmLinear	0.0817	0.0678	0.0645	0.075	1.09	0.0784	0.0002	0.0624	0.0773	1.013
svmRadial	0.075	0.0043	0.0596	0.075	1	0.0766	0.0325	0.0612	0.0773	0.990
nnet	0.0779	0.0374	0.0618	0.075	1.04	0.0815	0.0007	0.0633	0.0773	1.053
Unresidualized										
glmnet	0.0236	0.9043	0.0184	0.076	0.31	0.0265	0.8906	0.0209	0.0799	0.332
gbm	0.0256	0.8883	0.0201	0.076	0.34	0.0282	0.8772	0.0222	0.0799	0.352
svmLinear	0.0242	0.8998	0.0189	0.076	0.32	0.0272	0.8843	0.0215	0.0799	0.341
svmRadial	0.0248	0.8948	0.0193	0.076	0.32	0.0276	0.8813	0.0221	0.0799	0.345
nnet	1.4872	0.1117	1.4852	0.076	19.5	1.4870	0.0253	1.4849	0.0799	18.601

Note: RMSE = root mean square error. MAE = mean absolute error. SD = standard deviation.

Table 3.7 Model performance for brain structure predicting neuroticism and conscientiousness

	Training Data					Testing Data				
Algorithm	**RMSE**	**R^2**	**MAE**	**Outcome SD**	**RMSE/ SD**	**RMSE**	**R^2**	**MAE**	**Outcome SD**	**RMSE/ SD**
Neuroticism										
glmnet	22.701	0.005	17.98	22.72	0.999	21.592	0.00078	17.315	21.59	1.000
gbm	22.912	0.006	18.16	22.72	1.008	21.988	0.00446	17.592	21.59	1.018
svmLinear	25.818	0.006	20.69	22.72	1.136	23.156	0.00588	18.240	21.59	1.073
svmRadial	22.675	0.007	17.97	22.72	0.998	21.589	0.00001	17.268	21.59	1.000
nnet	22.674	0.007	17.95	22.72	0.998	21.569	0.00001	17.286	21.59	0.999
Conscientiousness										
glmnet	21.315	0.013	16.60	21.44	0.994	20.495	0.00019	16.136	20.41	1.004
gbm	21.351	0.016	16.72	21.44	0.996	20.778	0.00017	16.382	20.41	1.018
svmLinear	24.097	0.003	18.94	21.44	1.124	22.264	0.00017	17.568	20.41	1.091
svmRadial	21.330	0.015	16.61	21.44	0.995	20.514	0.00016	16.141	20.41	1.005
nnet	21.377	0.014	16.65	21.44	0.997	20.379	0.00078	16.058	20.41	0.999

Note: MRI features and personality scores were residualized. RMSE=root mean square error. MAE=mean absolute error. SD=standard deviation.

Chapter 4: Discussion

The current study aimed to develop and validate a predictive model of OCPD scores generated from normative personality data and to subsequently examine brain structure correlates of OCPD traits. Despite a moderate ability to predict OCPD traits using normative personality data that generalizes across samples, we found limited evidence that predicted OCPD scores are associated with individual differences in brain structure. Indeed, there was only one significant univariate association wherein thicker right superior frontal cortex was associated with higher OCPD scores. Adopting ML models to generate multivariate models of brain structure resulted in imprecise models and thus no reliable associations. Collectively, these data suggest that OCPD symptoms may be predicted using normative personality data, but that OCPD personality traits may not be strongly associated with brain structure and may require exceptionally large samples to reliably identify. Broadly, this approach exemplifies how deeply phenotyped small samples may be used to inform large national samples that may not have assessed specific phenotypes.

4.1 Predicting OCPD from the FFM

The first aim of this study was to develop and validate a predictive model of OCPD scores generated from normative personality data. Three OCPD measures (i.e., FFOCI, MAPP, and SIDP) were tested to determine which performed best. Broadly, the FFOCI models performed better than the MAPP and SIDP models. Normative personality data poorly predicted the total score of OCPD items (as a continuous variable) on the SIDP and MAPP, the two measures based on standard models of personality disorders. The limited number of individuals endorsing these items (n=46, 12 for baseline and FU10 SIDP OCPD diagnoses, respectively) may be contributing

to this poor performance, even though the models used a continuous variable of OCPD scores. This may also be due to the reduced overlap between DSM criteria-based variables and the FFM. In contrast, the FFOCI was developed from a dimensional model of PDs based on the FFM; of the three OCPD measures, the ML algorithms could best predict FFOCI scores from normative personality data. The larger, normal distribution of the FFOCI data may also have contributed to improved model performance over the SIDP and MAPP.

Evaluating machine learning performance metrics for dimensional outcomes is subjective and does not have the same established criteria as binary outcomes (e.g., sensitivity and specificity, area under the curve). The final trained OCPD model's RMSE (12.06 in the testing dataset) was approximately two-thirds of the FFOCI standard deviation. Performance was even better in the external validation dataset, suggesting results were not skewed by overfitting. The range of residualized FFOCI scores was roughly 80 points. Therefore, for the weighted average error between the predicted and actual scores to be 12 points seems adequate but not outstanding. These performance metrics are similar to another study employing machine learning to predict depression treatment outcomes from psychosocial predictor variables (RMSE/SD=0.79) (Webb et al., 2020). In previous research, meta-analyses of FFM models of OCPD (Samuel & Widiger, 2008; Saulsman & Page, 2004) found that the correlations between OCPD and the five factor model factors and facets ranged from -0.12 to 0.25. Although these are small effect sizes, this is typical in psychological research (Funder & Ozer, 2019; Götz et al., 2022). These small correlations may explain why the ML models predicted OCPD scores from personality data moderately well, and not higher.

When predicting OCPD scores from personality data, a distinct cluster of important NEO items, based on feature importance data, was not evident. This is not surprising given that the

model was trained on 114 individual NEO items. The NEO is designed to measure five personality factors, with six facets per factor and eight items per facet, so numerous items (i.e., predictor variables) were measuring similar constructs. The ten most important NEO items came from three out of the five factors (conscientiousness, openness, and extraversion), suggesting that OCPD is more than just conscientiousness. Interestingly, neuroticism was not even part of the top 15 items (**Figure 3.4**), while the other four personality factors were. This is likely due in part to fewer neuroticism items being kept in the model after feature selection; 18 of the 48 neuroticism items were included, while 30 conscientiousness, 26 openness, and 22 extraversion items were retained. Agreeableness also had 18 items in the final model and similarly had few items towards the top of feature importance. Additionally, neuroticism is only one-twelfth of the FFOCI-SF content, so it may be less critical to predicting FFOCI scores. Finally, another possible driver of neuroticism's limited presence amongst the most important features is the lower neuroticism scores compared to the other personality factors (e.g., median total score of ~70 vs. > 100). Thus, the neuroticism items may not load as strongly onto the OCPD outcome variables.

4.2 OCPD and Brain Structure

The second aim of this study was to investigate brain structure correlates of obsessive-compulsive personality traits. In standard regression analyses, where the predicted FFOCI score was the outcome variable and an individual brain region was the independent variable of interest, higher OCPD scores were significantly associated with thicker right superior frontal cortex. This finding is consistent with Payer and colleagues' (2015) results for individuals with any cluster C personality disorder symptoms. Right superior frontal gyrus has been linked to impulse control

(Hu et al., 2016). Thicker cortex here may contribute to the elevated levels of inhibitory control present in individuals with OCPD (American Psychiatric Association, 2013).

In the machine learning analyses, the neuroimaging models had large error metrics and were not able to accurately predict OCPD from brain structure. The best performing model, again elastic net regression, had a root mean square error equal to the outcome variable's standard deviation. As such, meaningful interpretations of the results cannot be drawn. Notably, another study employing machine learning to predict a continuous measure of a psychiatric diagnosis (i.e., depression severity) from brain structure obtained similar performance metrics (RMSE/SD ranged from 1.01 to 1.5 for various depression measures and samples) (Mwangi et al., 2012). It is noteworthy that three of the seven features retained in the final ML model were also identified in the standard regressions (right superior frontal thickness, left insula thickness, left fusiform surface area), although only one of these (right superior frontal thickness) was statistically significant after correcting for multiple comparisons.

There are two possible interpretations of the high ML model error. First, brain structure may truly not be associated with OCPD traits. If so, then it would make sense that the model cannot accurately predict OCPD scores from structural imaging data. The sample sizes of the sparse existing literature have been too small to inform this question. In general, large neuroimaging consortia have identified morphometric changes associated with many psychiatric disorders (Opel et al., 2020). However, effects are often small in magnitude (Marek et al., 2022). Personality has also been linked to brain structure (Hyatt et al., 2019; Owens et al., 2019; Riccelli et al., 2017) (Bjørnebekk et al., 2013; DeYoung et al., 2010; Privado et al., 2017; Vartanian et al., 2018), although there are also large studies that have found null, weak, or non-replicable results (Avinun et al., 2020; Gray et al., 2018; Hyatt et al., 2022; Masouleh et al., 2019; Nostro et al., 2017; Valk

et al., 2020). Of note, Avinun and colleagues analyzed the DNS dataset used in this study. If we are conceptualizing OCPD based on a five-factor model of personality, and personality was not significantly associated with brain structure in this dataset, then it may be unsurprising that OCPD is not associated with brain structure. However, those analyses were conducted using standard regression. As the authors suggested in their conclusion, it remained possible that multivariate machine learning approaches, which can account for variable interactions, linear and non-linear relationships, and both correlations between and uniqueness of variables, might identify brain structure correlates of personality. In the present study, post-hoc machine learning analyses for neuroticism and conscientiousness had model error as high as the OCPD analyses. Thus, the neuroimaging features did not reliably predict two key personality factors for OCPD.

The second possible interpretation is that model error was high due to limitations of the available data or error in the modeling techniques. To test for potential issues with the modeling techniques, a post-hoc analysis trained models to predict average cortical thickness from individual ROIs' thickness. The unresidualized models performed well, which decreases the likelihood that the results were due to error in the ML modeling process. The limitations of the available data are expanded upon below in the context of overall study limitations.

4.3 Limitations

First, I did not have access to a dataset with OCPD and neuroimaging data. It is unlikely for a neuroimaging sample that is large enough to conduct well-powered analyses to also have personality disorder assessment data. Large samples are needed given the expected small effects for associations between the brain and psychopathology or personality (Marek et al., 2022). To address this challenge, I drew from the literature on dimensional models of personality disorders.

This growing field of research has studied measurements of OCPD based on normative personality data and the Five-Factor model. This motivated me to employ machine learning techniques to predict OCPD scores from personality data, which is available in large neuroimaging datasets. The FFOCI measure was explicitly designed to capture OCPD from a dimensional, five factor model perspective. Perhaps unsurprisingly, the FFOCI was the most accurately predicted OCPD score from personality data. This approach permitted me to test associations between predicted OCPD scores and brain structure but did not allow for testing if findings for predicted scores were the same as for observed scores.

Second, there were demographic differences between the sample used for the OCPD model training (SPAN) and the sample that the model was applied to (DNS). To address this, I used a third dataset with age and education distributions similar to DNS to assess the external validity of the OCPD model. Model performance was similar in the SPAN testing and FFOCI external validation datasets. In fact, the error was smaller in the external validation dataset (RMSE/SD = 0.51 for residualized, compared to 0.66 in SPAN). This was important given that model training occurred in an older middle-aged sample (SPAN) yet the model was generated to be applied to a college student sample (DNS). The FFOCI validation study, which was also a college student sample, allowed us to test if our model would be valid on the younger DNS sample. Residualizing the data to control for age and gender generated the most accurate predictions. Therefore, the residualized version of the model was used to predict OCPD scores in the DNS dataset.

Third, the OCPD model performance was moderate, limiting the accuracy of predicted OCPD scores and informativeness of correlations with other phenotypes such as brain structure. This could be due to the likely small sample size of individuals meeting criteria for OCPD at timepoint 12, the timepoint with the best performing model. The SIDP was not administered at

that visit, but at timepoint 10 only 12 people met criteria according to the clinician-administered interview. On the MAPP self-report personality disorder assessment, the number of participants meeting diagnostic criteria decreased from 57 at timepoint 10 (notably much higher than 12 as identified by the SIDP) to 49 at timepoint 12. Therefore, it is likely that the SIDP numbers would have also decreased. Unfortunately, the FFOCI was not developed at the time that the SPAN study began, so baseline FFOCI data, when the number of OCPD diagnoses was highest (n=48), is not available. Model performance for the other two OCPD measures available at baseline, SIDP and MAPP, was much worse than the FFOCI. This is possibly due to smaller score distributions of these measures and/or their DSM diagnostic criteria focus rather than a dimensional model. The trained OCPD model performed better in the external validation dataset, which had been oversampled for OCPD, compared to the SPAN testing dataset. However, model performance in the SPAN validation dataset, used to select the final OCPD model, was similar to the external dataset (SPAN RMSE/SD=0.52, external RMSE/SD=0.51). These external validation results suggest that the OCPD model generalizes and is not overfitting.

It remains possible that the ML neuroimaging models were inaccurate because the predicted OCPD scores that were generated in Aim 1 and used as the training label (i.e., outcome variable) were imprecise. Unfortunately, this cannot be directly tested as I do not have a neuroimaging dataset with observed OCPD data. Like in the OCPD model, the small number of individuals meeting diagnostic criteria for OCPD in the Aim 1 training dataset could be causing the high ML neuroimaging model error. There may not be enough OCPD "signal" to properly train the ML algorithms. Interestingly, median clinician-rated scores of OCPD, based on SIDP total scores for OCPD questions, decreased between baseline and timepoint 10, while self-report MAPP

scores stayed the same. Given our dimensional conceptualization of OCPD, we did not believe the small sample size meeting strict diagnostic criteria would invalidate the study.

4.4 Implications and Future Directions

The present study is situated in the broader context of dimensional models of personality disorders, OCPD assessment, machine learning-based prediction, and brain structure correlates of psychiatric disorders. The results suggest that normative personality data can reasonably predict an FFM-based personality disorder measure, although performance for traditional assessments (i.e., SPAN, MAPP) is weak. Future studies are needed to investigate whether this is an inherent limitation of traditional OCPD classification or potentially a result of the present study's small diagnostic sample size. Predicted OCPD trait scores had similar expected correlations with depression and marital satisfaction as the observed OCPD trait scores. Additionally, OCPD traits may be associated with thicker right superior frontal cortex, but the current study was unable to detect significant brain structure associations using machine learning techniques. Future research with a sufficient sample size of neuroimaging and OCPD data is needed to determine whether brain structure is indeed associated with OCPD. Until this has been conducted, using the current study's approach for other personality disorders is unlikely to generate results from which the field can draw reliable conclusions. A focused medium-sized OCPD sample may provide the evidence needed to support employing this approach in large epidemiological neuroimaging studies for the full array of personality disorders.

References

Abi-Dargham, A., & Horga, G. (2016). The search for imaging biomarkers in psychiatric disorders. *Nature Medicine 2016 22:11*, *22*(11), 1248–1255. https://doi.org/10.1038/nm.4190

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edit). https://doi.org/10.1176/appi.books.9780890425596.744053

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences*, *18*(4), 177–185. https://doi.org/10.1016/j.tics.2013.12.003

Atmaca, M., Korucu, T., Caglar Kilic, M., Kazgan, A., & Yildirim, H. (2019). Pineal gland volumes are changed in patients with obsessive-compulsive personality disorder. *Journal of Clinical Neuroscience: Official Journal of the Neurosurgical Society of Australasia*, *70*, 221–225. https://doi.org/10.1016/j.jocn.2019.07.047

Atmaca, M., Korucu, T., Tabara, M. F., Yildirim, H., & Kılıç, M. C. (2019). Volumetric MRI study of orbito-frontal cortex and thalamus in obsessive-compulsive personality disorder. *Journal of Clinical Neuroscience: Official Journal of the Neurosurgical Society of Australasia*, *64*, 89–93. https://doi.org/10.1016/j.jocn.2019.03.062

Atroszko, P. A., Demetrovics, Z., & Griffiths, M. D. (2020). Work addiction, obsessive-compulsive personality disorder, burn-out, and global burden of disease: Implications from the ICD-11. *International Journal of Environmental Research and Public Health*, *17*(2). https://doi.org/10.3390/ijerph17020660

Avinun, R., Israel, S., Knodt, A. R., & Hariri, A. R. (2020). Little evidence for associations between the Big Five personality traits and variability in brain gray or white matter. *NeuroImage*, *220*, 117092. https://doi.org/10.1016/j.neuroimage.2020.117092

Baranger, D. A. A., Few, L. R., Sheinbein, D. H., Agrawal, A., Oltmanns, T. F., Knodt, A. R., Barch, D. M., Hariri, A. R., & Bogdan, R. (2020). Borderline Personality Traits Are Not Correlated With Brain Structure in Two Large Samples. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(7), 669–677. https://doi.org/10.1016/j.bpsc.2020.02.006

Bechara A., Damasio H., & Damasio A. (2000). Emotion, Decision Making and the Orbitofrontal Cortex. *Cerebral Cortex*, *10*, 295–307. internal-pdf://emot_decision_bechara_damasio-1566890261/emot_decision_bechara_damasio.pdf

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory (BDI-II): Manual and Questionnaire*. The Psychological Corporation.

Beeney, J. E., Forbes, E. E., Hipwell, A. E., Nance, M., Mattia, A., Lawless, J. M., Banihashemi, L., & Stepp, S. D. (2021). Determining the key childhood and adolescent risk factors for future BPD symptoms using regularized regression: comparison to depression and conduct disorder. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *62*(2), 223–231. https://doi.org/10.1111/jcpp.13269

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. https://doi.org/10.1016/j.neuroimage.2012.10.009

Brewer, J. A., Garrison, K. A., & Whitfield-Gabrieli, S. (2013). What about the "self" is processed in the posterior cingulate cortex? *Frontiers in Human Neuroscience*, *7*(OCT), 647. https://doi.org/10.3389/fnhum.2013.00647

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Cardinal, R. N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. In *Neuroscience and Biobehavioral Reviews* (Vol. 26, Issue 3, pp. 321–352). Pergamon. https://doi.org/10.1016/S0149-7634(02)00007-6

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. In *Cognitive, Affective and Behavioral Neuroscience* (Vol. 7, Issue 4, pp. 367–379). Springer. https://doi.org/10.3758/CABN.7.4.367

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. In *Brain* (Vol. 129, Issue 3, pp. 564–583). Oxford University Press. https://doi.org/10.1093/brain/awl004

Clark, L. A., Livesley, W. J., Schroeder, M. L., & Irish, S. L. (1996). Convergence of two systems for assessing specific trait disorders. *Psychological Assessment*, *8*(3), 294–303.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI–R) and NEO Five-Factor Inventory (NEO–FFI) Professional Manual*. Psychological Assessment Resources.

Coutinho, J., Goncalves, O. F., Soares, J. M., Marques, P., & Sampaio, A. (2016). Alterations of the default mode network connectivity in obsessive–compulsive personality disorder: A pilot study. *Psychiatry Research - Neuroimaging*, *256*, 1–7. https://doi.org/10.1016/j.pscychresns.2016.08.007

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

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*(6), 820–828. https://doi.org/10.1177/0956797610370159

Dick, A. S., Lopez, D. A., Watts, A. L., Heeringa, S., Reuter, C., Bartsch, H., Fan, C. C., Kennedy, D. N., Palmer, C., Marshall, A., Haist, F., Hawes, S., Nichols, T. E., Barch, D. M., Jernigan, T. L., Garavan, H., Grant, S., Pariyadath, V., Hoffman, E., … Thompson, W. K. (2021). Meaningful associations in the adolescent brain cognitive development study. *NeuroImage*, *239*(June), 118262. https://doi.org/10.1016/j.neuroimage.2021.118262

Dick, A. S., Watts, A. L., Heeringa, S., Lopez, D. A., Bartsch, H., Chieh Fan, C., Palmer, C., Reuter, C., Marshall, A., Haist, F., Hawes, S., Nichols, T. E., Barch, D. M., Jernigan, T. L., Garavan, H., Grant, S., Pariyadath, V., Hoffman, E., Neale, M., … Thompson, W. K. (2020). Meaningful Effects in the Adolescent Brain Cognitive Development Study. *BioRxiv*, 2020.09.01.276451. https://doi.org/10.1101/2020.09.01.276451

Diener, E., Emmon, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75.

Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, *41*, 417–440. https://doi.org/10.1007/978-3-319-24612-3_300988

Drucker, H., Surges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support vector regression machines. *Advances in Neural Information Processing Systems*, *1*, 155–161.
Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T. E., Caspi, A., & Hariri, A. R. (2020). What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis. *Psychological Science*, *31*(7), 792–806. https://doi.org/10.1177/0956797620916786
Fernandes Gonçalves, M. (2015). *Structural Connectivity of the Default Mode Network in Obsessive-Compulsive Personality Disorder*. Universidade do Minho.
Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341–355. https://doi.org/10.1016/S0896-6273(02)00569-X
Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195–207. https://doi.org/10.1006/nimg.1998.0396
Friedman, J. (2001). Greedy Function Approximation: A Gradient Boosting Machine P. *The Annals of Statistics*, *29*(5), 1189–1232.
Friedman, N. P., & Robbins, T. W. (2022). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*, *47*(1), 72–89. https://doi.org/10.1038/s41386-021-01132-0
Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202
García-Villamisar, D., & Dattilo, J. (2015). Executive functioning in people with obsessive-compulsive personality traits: Evidence of modest impairment. *Journal of Personality Disorders*, *29*(3), 418–430. https://doi.org/10.1521/pedi_2013_27_101
Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., Hampton, J. M., Coalson, R. S., Nguyen, A. L., McDermott, K. B., Shimony, J. S., Snyder, A. Z., Schlaggar, B. L., Petersen, S. E., Nelson, S. M., & Dosenbach, N. U. F. (2017). Precision Functional Mapping of Individual Human Brains. *Neuron*, *95*(4), 791-807.e7. https://doi.org/10.1016/j.neuron.2017.07.011
Götz, F. M., Gosling, S. D., & Rentfrow, P. J. (2022). Small Effects: The Indispensable Foundation for a Cumulative Psychological Science. *Perspectives on Psychological Science*, *17*(1), 205–215. https://doi.org/10.1177/1745691620984483
Grahn, J. A., Parkinson, J. A., & Owen, A. M. (2008). The cognitive functions of the caudate nucleus. *Progress in Neurobiology*, *86*(3), 141–155. https://doi.org/10.1016/j.pneurobio.2008.09.004
Grant, J. E., Mooney, M. E., & Kushner, M. G. (2012). Prevalence, correlates, and comorbidity of DSM-IV obsessive-compulsive personality disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Psychiatric Research*, *46*(4), 469–475. https://doi.org/10.1016/j.jpsychires.2012.01.009
Gray, J. C., Owens, M. M., Hyatt, C. S., & Miller, J. D. (2018). No evidence for morphometric associations of the amygdala and hippocampus with the five-factor model personality traits in relatively healthy young adults. *PLoS ONE*, *13*(9), e0204011. https://doi.org/10.1371/journal.pone.0204011

Griffin, S. A., Suzuki, T., Lynam, D. R., Crego, C., Widiger, T. A., Miller, J. D., & Samuel, D. B. (2018). Development and Examination of the Five-Factor Obsessive-Compulsive Inventory-Short Form. *Assessment*, *25*(1), 56–68. https://doi.org/10.1177/1073191116643818

Grzenda, A., Kraguljac, N. v., McDonald, W. M., Nemeroff, C., Torous, J., Alpert, J. E., Rodriguez, C. I., & Widge, A. S. (2021). Evaluating the Machine Learning Literature: A Primer and User's Guide for Psychiatrists. *American Journal of Psychiatry*, 1–14.

Gurok, M. G., Korucu, T., Kilic, M. C., Yildirim, H., & Atmaca, M. (2019). Hippocampus and amygdalar volumes in patients with obsessive-compulsive personality disorder. *Journal of Clinical Neuroscience*, *64*, 259–263. https://doi.org/10.1016/j.jocn.2019.03.060

Haas, B. W., & Miller, J. D. (2015). Borderline Personality Traits and Brain Activity During Emotional Perspective Taking. *Personality Disorders: Theory, Research, and Treatment*, *6*(4), 315–320. http://dx.doi.org/10.1037/per0000130%5Cnhttp://dx.doi.org/10.1037/per0000130.supp

Haigler, E. D., & Widiger, T. A. (2001). Experimental manipulation of NEO-PI-R items. *Journal of Personality Assessment*, *77*(2), 339–358. https://doi.org/10.1207/S15327752JPA7702_14

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., Maguire, P., Rosas, D., Makris, N., Dale, A., Dickerson, B., & Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, *32*(1), 180–194. https://doi.org/10.1016/j.neuroimage.2006.02.051

Horvath, A., Dras, M., Lai, C. C. W., & Boag, S. (2020). Predicting Suicidal Behavior Without Asking About Suicidal Ideation: Machine Learning and the Role of Borderline Personality Disorder Criteria. *Suicide and Life-Threatening Behavior*, *July*, 1–12. https://doi.org/10.1111/sltb.12719

Hu, S., Ide, J. S., Zhang, S., & Li, C. S. R. (2016). The right superior frontal gyrus and individual variation in proactive control of impulsive response. *Journal of Neuroscience*, *36*(50), 12688–12696. https://doi.org/10.1523/JNEUROSCI.1175-16.2016

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), 1–11. https://doi.org/10.1037/abn0000391

Hyatt, C. S., Sharpe, B. M., Owens, M. M., Listyg, B. S., Carter, N. T., Lynam, D. R., & Miller, J. D. (2022). Searching High and Low for Meaningful and Replicable Morphometric Correlates of Personality. *Journal of Personality and Social Psychology*, *123*(2), 463–480. https://doi.org/10.1037/pspp0000402

Kaller, C. P., Rahm, B., Spreer, J., Weiller, C., & Unterrainer, J. M. (2011). Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. *Cerebral Cortex*, *21*(2), 307–317. https://doi.org/10.1093/cercor/bhq096

Klein, A., & Tourville, J. (2012). 101 labeled brain images and a consistent human cortical labeling protocol. *Frontiers in Neuroscience*, *0*(DEC), 171. https://doi.org/10.3389/fnins.2012.00171

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Lawton, E. M. (2014). *Reconceptualizing social impairment using informant report*.

Lei, H., Huang, L., Li, J., Liu, W., Fan, J., Zhang, X., Xia, J., Zhao, K., Zhu, X., & Rao, H. (2020). Altered spontaneous brain activity in obsessive-compulsive personality disorder. *Comprehensive Psychiatry*, *96*, 152144. https://doi.org/10.1016/j.comppsych.2019.152144

Li, J., & Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, *95*(3), 221–227. https://doi.org/10.1038/sj.hdy.6800717

Luo, Y., Chen, L., Li, H., Dong, Y., Zhou, X., Qiu, L., Zhang, L., Gao, Y., Zhu, C., Yu, F., & Wang, K. (2020). Do Individuals With Obsessive-Compulsive Disorder and Obsessive-Compulsive Personality Disorder Share Similar Neural Mechanisms of Decision-Making Under Ambiguous Circumstances? *Frontiers in Human Neuroscience*, *14*, 441. https://doi.org/10.3389/fnhum.2020.585086

Lynam, D. R., & Widiger, T. A. (2001). Using the Five-Factor Model to Represent the DSM-IV Personality Disorders: An Expert Consensus Approach. *Journal of Abnormal Psychology*, *110*(3), 401–412. https://doi.org/10.1037//0021-843X.

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., … Dosenbach, N. U. F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*(7902), 654–660. https://doi.org/10.1038/s41586-022-04492-9

Marincowitz, C., Lochner, C., & Stein, D. J. (2021). The neurobiology of obsessive-compulsive personality disorder. *CNS Spectrums*, 1–12. https://doi.org/https://doi.org/10.1017/S1092852921000754

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, *88*(1), 139–157. https://doi.org/10.1037/0022-3514.88.1.139

Masouleh, S. K., Eickhoff, S. B., Hoffstaedter, F., & Genon, S. (2019). Empirical examination of the replicability of associations between brain structure and psychological variables. *ELife*, *8*, 1–25. https://doi.org/10.7554/eLife.43464

Mike, A., King, H., Oltmanns, T. F., & Jackson, J. J. (2018). Obsessive, compulsive, and conscientious? The relationship between OCPD and personality traits. *Journal of Personality*, *86*(6), 952–972. https://doi.org/10.1111/jopy.12368

Morris, S. E., & Cuthbert, B. N. (2012). Research Domain Criteria: cognitive systems, neural circuits, and dimensions of behavior. *Dialogues in Clinical Neuroscience*, *14*(1), 29–37. https://doi.org/10.1016/0015-1882(90)80068-v

Mushiake, H., Saito, N., Sakamoto, K., Itoyama, Y., & Tanji, J. (2006). Activity in the Lateral Prefrontal Cortex Reflects Multiple Steps of Future Events in Action Plans. *Neuron*, *50*(4), 631–641. https://doi.org/10.1016/j.neuron.2006.03.045

Mutschler, I., Wieckhorst, B., Kowalevski, S., Derix, J., Wentlandt, J., Schulze-Bonhage, A., & Ball, T. (2009). Functional organization of the human anterior insular cortex. *Neuroscience Letters*, *457*(2), 66–70. https://doi.org/10.1016/j.neulet.2009.03.101

Mwangi, B., Matthews, K., & Steele, J. D. (2012). Prediction of illness severity in patients with major depression using structural MR brain scans. *Journal of Magnetic Resonance Imaging*, *35*(1), 64–71. https://doi.org/10.1002/jmri.22806

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, *7*(DEC), 21. https://doi.org/10.3389/fnbot.2013.00021

Nielsen, A. N., Barch, D. M., Petersen, S. E., Schlaggar, B. L., & Greene, D. J. (2020). Machine Learning With Neuroimaging: Evaluating Its Applications in Psychiatry. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(8), 791–798. https://doi.org/10.1016/j.bpsc.2019.11.007

Nikolova, Y. S., Bogdan, R., Brigidi, B. D., & Hariri, A. R. (2012). Ventral Striatum Reactivity to Reward and Recent Life Stress Interact to Predict Positive Affect. *Biological Psychiatry*, *72*(2), 157–163. https://doi.org/10.1016/J.BIOPSYCH.2012.03.014

Nostro, A. D., Müller, V. I., Reid, A. T., & Eickhoff, S. B. (2017). Correlations between Personality and Brain Structure: A Crucial Role of Gender. *Cerebral Cortex*, *27*(7), 3698–3712. https://doi.org/10.1093/cercor/bhw191

Nyholt, D. R. (2004). A Simple Correction for Multiple Testing for Single-Nucleotide Polymorphisms in Linkage Disequilibrium with Each Other. *The American Journal of Human Genetics*, *74*(4), 765–769. https://doi.org/10.1086/383251

O'Connor, B. P. (2005). A search for consensus on the dimensional structure of personality disorders. *Journal of Clinical Psychology*, *61*(3), 323–345. https://doi.org/10.1002/jclp.20017

Oltmanns, T. F., & Gleason, M. E. J. (2011). Personality Pathology, Health, and Social Adjustment in Later Life. *Mental Health in Public Health: The Next 100 Years*, *December*, 151–179.

Oltmanns, T. F., Rodrigues, M. M., Weinstein, Y., & Gleason, M. E. J. (2014). Prevalence of Personality Disorders at Midlife in a Community Sample: Disorders and Symptoms Reflected in Interview, Self, and Informant Reports. *Journal of Psychopathology and Behavioral Assessment*, *36*(2), 177–188. https://doi.org/10.1007/s10862-013-9389-7

Oltmanns, T. F., Turkheimer, E., & Strauss, M. E. (1998). Peer assessment of personality traits and pathology in female college students. *Assessment*, *5*(1), 53–65. https://doi.org/10.1177/107319119800500108

Opel, N., Goltermann, J., Hermesdorf, M., Berger, K., Baune, B. T., & Dannlowski, U. (2020). Cross-disorder analysis of brain structural abnormalities in six major psychiatric disorders: A secondary analysis of mega- and meta-analytical findings from the ENIGMA consortium. *Biological Psychiatry*, *88*(9), 678–686. https://doi.org/10.1016/j.biopsych.2020.04.027

Owens, M. M., Hyatt, C. S., Gray, J. C., Carter, N. T., Mackillop, J., Miller, J. D., & Sweet, L. H. (2019). Cortical morphometry of the five-factor model of personality: Findings from the Human Connectome Project full sample. *Social Cognitive and Affective Neuroscience*, *14*(4), 381–395. https://doi.org/10.1093/scan/nsz017

Payer, D. E., Park, M. T. M., Kish, S. J., Kolla, N. J., Lerch, J. P., Boileau, I., & Chakravarty, M. M. (2015). Personality disorder symptomatology is associated with anomalies in striatal and prefrontal morphology. *Frontiers in Human Neuroscience*, *9*(August). https://doi.org/10.3389/fnhum.2015.00472

Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured Interview for DSM-IV Personality: SIDP-IV*. American Psychiatric Press. https://books.google.com/books?id=J17zkm2RH6MC

Pierce, G. R., Sarason, I. G., & Sarason, B. R. (1991). General and relationship-based perceptions of social support. *Journal of Personality and Social Psychology*, *61*(6), 1028–1039. http://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=1992-12680-001&S=L&D=pdh&EbscoContent=dGJyMNXb4kSep7c4yOvsOLCmr1Gep7VSrqu4TLCWxWXS&ContentCustomer=dGJyMPGrtE%2Bwp7dMuePfgeyx44Dt6fIA

Privado, J., Román, F. J., Saénz-Urturi, C., Burgaleta, M., & Colom, R. (2017). Gray and white matter correlates of the Big Five personality traits. *Neuroscience*, *349*, 174–184. https://doi.org/10.1016/j.neuroscience.2017.02.039

Raichle, M. E. (2015). *The Brain's Default Mode Network*. https://doi.org/10.1146/annurev-neuro-071013-014030

Rawls, E., Miskovic, V., Moody, S. N., Lee, Y., Shirtcliff, E. A., & Lamm, C. (2020). Feedback-Related Negativity and Frontal Midline Theta Reflect Dissociable Processing of Reinforcement. *Frontiers in Human Neuroscience*, *13*(January), 1–14. https://doi.org/10.3389/fnhum.2019.00452

Reddan, M. C., Lindquist, M. A., & Wager, T. D. (2017). Effect size estimation in neuroimaging. In *JAMA Psychiatry* (Vol. 74, Issue 3, pp. 207–208). American Medical Association. https://doi.org/10.1001/jamapsychiatry.2016.3356

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*(4), 671–684. https://doi.org/10.1093/scan/nsw175

Richter, T., Fishbain, B., Markus, A., Richter-Levin, G., & Okon-Singer, H. (2020). Using machine learning-based analysis for behavioral differentiation between anxiety and depression. *Scientific Reports*, *10*(1), 16381. https://doi.org/10.1038/s41598-020-72289-9

Russell, D. W. (1996). UCLA Loneliness Scale (Version 3): Reliability, Validity, and Factor Structure. *Journal of Personality Assessment*, *66(1)*(Version 3), 20–40.

Sabourin, S., Valois, P., & Lussier, Y. (2005). Development and Validation of a Brief Version of the Dyadic Adjustment Scale with a Nonparametric Item Analysis Model. *Psychological Assessment*, *17*(1), 15–27.

Samuel, D. B., Balling, C. E., & Bucher, M. A. (2022). The alternative model of personality disorder is inadequate for capturing obsessive-compulsive personality disorder. *Personality Disorders: Theory, Research, and Treatment*, *13*(4), 418–421. https://doi.org/10.1037/per0000544

Samuel, D. B., Riddell, A. D. B., Lynam, D. R., Miller, J. D., & Widiger, T. A. (2012). A five-factor measure of obsessive-compulsive personality traits. *Journal of Personality Assessment*, *94*(5), 456–465. https://doi.org/10.1080/00223891.2012.677885

Samuel, D. B., & Widiger, T. A. (2004). Clinicians' personality descriptions of prototypic personality disorders. *Journal of Personality Disorders*, *18*(3), 286–308. https://doi.org/10.1521/pedi.18.3.286.35446

Samuel, D. B., & Widiger, T. A. (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: A facet level analysis. *Clinical Psychology Review*, *28*(8), 1326–1342. https://doi.org/10.1016/j.cpr.2008.07.002

Samuel, D. B., & Widiger, T. A. (2010). A Comparison of Obsessive-Compulsive Personality Disorder scales. *Journal of Personality Assessment*, *92*(3), 232–240. https://doi.org/10.1080/00223891003670182

Samuel, D. B., & Widiger, T. A. (2011). Conscientiousness and obsessive-compulsive personality disorder. *Personality Disorders: Theory, Research, and Treatment*, *2*(3), 161–174. https://doi.org/10.1037/a0021216

Saulsman, L. M., & Page, A. C. (2004). The five-factor model and personality disorder empirical literature: A meta-analytic review. *Clinical Psychology Review*, *23*(8), 1055–1085. https://doi.org/10.1016/j.cpr.2002.09.001

Schroeder, M. L., Wormworth, J. A., & Livesley, W. J. (1992). Dimensions of Personality Disorder and Their Relationships to the Big Five Dimensions of Personality. *Psychological Assessment*, *4*(1), 47–53. https://doi.org/10.1037/1040-3590.4.1.47

Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, *49*(9), 1426–1448. https://doi.org/10.1017/S0033291719000151

Shenhav, A., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. In *Nature Neuroscience* (Vol. 19, Issue 10, pp. 1286–1291). Nature Publishing Group. https://doi.org/10.1038/nn.4384

Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, *13*(8), 334–340. https://doi.org/10.1016/j.tics.2009.05.001

Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Dyck, I. R., Stout, R. L., Bender, D. S., Grilo, C. M., Tracie Shea, M., Zanarini, M. C., Morey, L. C., Sanislow, C. A., & Oldham, J. M. (2002). Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *American Journal of Psychiatry*, *159*(2), 276–283. https://doi.org/10.1176/appi.ajp.159.2.276

Skodol, A. E., Pagano, M. E., Bender, D. S., Shea, M. T., Gunderson, J. G., Yen, S., Stout, R. L., Morey, L. C., Sanislow, C. A., Grilo, C. M., Zanarini, M. C., & Mcglashan, T. H. (2005). Stability of functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personailty disorder over two years. *Psychological Medicine*, *35*(3), 443–451. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf

Smith, J. M., Grandin, L. D., Alloy, L. B., & Abramson, L. Y. (2006). Cognitive vulnerability to depression and Axis II personality dysfunction. *Cognitive Therapy and Research*, *30*(5), 609–621. https://doi.org/10.1007/s10608-006-9038-5

Soares, J. M., Marques, P., Alves, V., & Sousa, N. (2013). A hitchhiker's guide to diffusion tensor imaging. *Frontiers in Neuroscience*, *7*(31), 1–14. https://doi.org/10.3389/fnins.2013.00031

Soeteman, D. I., Roijen, L. H. van, Verheul, R., & Busschbach, J. J. V. (2008). The economic burden of personality disorders in mental health care. *Journal of Clinical Psychiatry*, *69*(2), 259–265. https://doi.org/10.4088/JCP.v69n0212

South, S. C., Boudreaux, M. J., & Oltmanns, T. F. (2020). The impact of personality disorders on longitudinal change in relationship satisfaction in long-term married couples. *Journal of Personality Disorders*, *34*(4), 439–458. https://doi.org/10.1521/pedi_2019_33_392

Tang, W., Jbabdi, S., Zhu, Z., Cottaar, M., Grisot, G., Lehman, J. F., Yendiki, A., & Haber, S. N. (2019). A connectional hub in the rostral anterior cingulate cortex links areas of emotion and cognitive control. *ELIFE*, *8*. https://doi.org/10.7554/eLife.43761

Tanga, Y., Hojatkashanib, C., Dinovb, I. D., Suna, B., Fana, L., Lina, X., Qic, H., Huab, X., Liua, S., & Toga, A. W. (2010). A morphometric comparison study between Chinese and Caucasian cohorts. *NeuroImage*, *51*(1), 33–41. https://doi.org/10.1016/j.neuroimage.2010.01.111.The

Valk, S. L., Hoffstaedter, F., Camilleri, J. A., Kochunov, P., Yeo, B. T. T., & Eickhoff, S. B. (2020). Personality and local brain structure: Their shared genetic basis and reproducibility. *NeuroImage*, *220*, 117067. https://doi.org/10.1016/j.neuroimage.2020.117067

Vartanian, O., Wertz, C. J., Flores, R. A., Beatty, E. L., Smith, I., Blackler, K., Lam, Q., & Jung, R. E. (2018). Structural correlates of Openness and Intellect: Implications for the

contribution of personality to creativity. *Human Brain Mapping*, *39*(7), 2987–2996. https://doi.org/10.1002/hbm.24054

Vickers, R. R., Conway, T. L., & Hervig, L. K. (1990). Demonstration of Replicable Dimensions of Health Behaviors. *Preventive Medicine*, *19*(4), 377–401.

Volkert, J., Gablonski, T.-C., & Rabung, S. (2018). Prevalence of personality disorders in the general adult population in Western countries: Systematic review and meta-analysis. *The British Journal of Psychiatry*, *213*(6), 709–715. https://doi.org/10.1192/bjp.2018.202

Walter, M., Alizadeh, S., Jamalabadi, H., Lueken, U., Dannlowski, U., Walter, H., Olbrich, S., Colic, L., Kambeitz, J., Koutsouleris, N., Hahn, T., & Dwyer, D. B. (2019a). Translational machine learning for psychiatric neuroimaging. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *91*, 113–121. https://doi.org/10.1016/j.pnpbp.2018.09.014

Walter, M., Alizadeh, S., Jamalabadi, H., Lueken, U., Dannlowski, U., Walter, H., Olbrich, S., Colic, L., Kambeitz, J., Koutsouleris, N., Hahn, T., & Dwyer, D. B. (2019b). Translational machine learning for psychiatric neuroimaging. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *91*, 113–121. https://doi.org/10.1016/j.pnpbp.2018.09.014

Walton, M. E., Behrens, T. E. J., Noonan, M. P., & Rushworth, M. F. S. (2011). Giving credit where credit is due: Orbitofrontal cortex and valuation in an uncertain world. *Annals of the New York Academy of Sciences*, *1239*(1), 14–24. https://doi.org/10.1111/j.1749-6632.2011.06257.x

Watson, D., Clark, L. A., & Chmielewski, M. (2008). Structures of personality and their relevance to psychopathology: II. Further articulation of a comprehensive unified trait structure. *Journal of Personality*, *76*(6), 1545–1586. https://doi.org/10.1111/j.1467-6494.2008.00531.x

Webb, C. A., Cohen, Z. D., Beard, C., Forgeard, M., Peckham, A. D., Björgvinsson, T., & Al, W. E. T. (2020). Personalized Prognostic Prediction of Treatment Outcome for Depressed Patients in a Naturalistic Psychiatric Hospital Setting: A Comparison of Machine Learning Approaches. *Journal of Consulting and Clinical Psychology*, *88*(1), 25–38. https://doi.org/10.1037/ccp0000451.supp

Weissman, M. M. (1999). *SAS-SR Question Booklet*. Multi Health Systems.

Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *American Psychologist*, *62*(2), 71–83. https://doi.org/10.1037/0003-066X.62.2.71

Zhou, H. X., Chen, X., Shen, Y. Q., Li, L., Chen, N. X., Zhu, Z. C., Castellanos, F. X., & Yan, C. G. (2020). Rumination and the default mode network: Meta-analysis of brain imaging studies and implications for depression. In *NeuroImage* (Vol. 206, p. 116287). Academic Press Inc. https://doi.org/10.1016/j.neuroimage.2019.116287

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x