Schizophrenia Bulletin
https://doi.org/10.1093/schbul/sbac216

Multivariate Associations Among White Matter, Neurocognition, and Social Cognition Across Individuals With Schizophrenia Spectrum Disorders and Healthy Controls

Navona Calarco[1,*], Lindsay D. Oliver[1], Michael Joseph[1], Colin Hawco[1,2], Erin W. Dickie[1], Pamela DeRosse[3,4,5], James M. Gold[6], George Foussias[1,2], Miklos Argyelan[3,4,5], Anil K. Malhotra[3,4,5], Robert W. Buchanan[6], and Aristotle N. Voineskos[*,1,2,*] for the SPINS Group

[1]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada; [2]Department of Psychiatry, University of Toronto, Toronto, ON, Canada; [3]Division of Psychiatry Research, Division of Northwell Health, The Zucker Hillside Hospital, Glen Oaks, NY, USA; [4]Department of Psychiatry, The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, USA; [5]Center for Psychiatric Neuroscience, The Feinstein Institute for Medical Research, Manhasset, NY, USA; [6]Department of Psychiatry, Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore, MD, USA

[*]To whom correspondence should be addressed; Centre for Addiction and Mental Health, 250 College St., Toronto, ON, Canada, M5T 1R8; tel: 416-535-8501 x34378, e-mail: aristotle.voineskos@camh.ca

***Background and Hypothesis*: Neurocognitive and social cognitive abilities are important contributors to functional outcomes in schizophrenia spectrum disorders (SSDs). An unanswered question of considerable interest is whether neurocognitive and social cognitive deficits arise from overlapping or distinct white matter impairment(s). *Study Design*: We sought to fill this gap, by harnessing a large sample of individuals from the multi-center Social Processes Initiative in the Neurobiology of the Schizophrenia(s) (SPINS) dataset, unique in its collection of advanced diffusion imaging and an extensive battery of cognitive assessments. We applied canonical correlation analysis to estimates of white matter microstructure, and cognitive performance, across people with and without an SSD. *Study Results*: Our results established that white matter circuitry is dimensionally and strongly related to both neurocognition and social cognition, and that microstructure of the uncinate fasciculus and the rostral body of the corpus callosum may assume a "privileged role" subserving both. Further, we found that participant-wise estimates of white matter microstructure, weighted by cognitive performance, were largely consistent with participants' categorical diagnosis, and predictive of (cross-sectional) functional outcomes. *Conclusions*: The demonstrated strength of the relationship between white matter circuitry and neurocognition and social cognition underscores the potential for using relationships among these variables to identify biomarkers of functioning, with potential prognostic and therapeutic implications.**

Key words: schizophrenia spectrum disorders/neurocognition/social cognition/diffusion imaging/white matter/Research Domain Criteria

Introduction

Neurocognitive and social cognitive deficits are pervasive in schizophrenia spectrum disorders (SSD),[1,2] and tightly bound to poor functional outcomes.[3–5] Several lines of evidence have established that neurocognition and social cognition are behaviorally separable[3,6] but an open question, relevant for treatment discovery efforts, is if neurocognition and social cognition share a common biological basis.

Some recent research has approached this question using functional MRI. For example, our group has shown that neurocognitive and social cognitive performance in SSD can be predicted by resting-state connectivity in the mirror neuron and mentalizing systems, and that connectivity is predictive of functional outcomes.[7] To the best of our knowledge, no study has taken a comparable approach in white matter. This is important, because associations have been shown between white matter structure and social cognition, and white matter structure and neurocognition. However, social cognition and neurocognition are correlated[8]; therefore it is possible that they may share similar neural underpinnings, but some might be distinct. Finally, diffusion MRI has potential as a tool for clinical translation if there is interest

ultimately in early biological prognostic indicators of social cognition and neurocognition, as well as targeting for therapeutics. In particular, advances in diffusion MRI now allow high quality sequence acquisition in short periods of time.[9]

Despite the paucity of studies examining the cognitions in tandem, a large number of neuroimaging investigations have established that disruption of white matter is associated with neurocognitive deficits, demonstrating that several long-range, deep white matter tracts are reliably impaired in SSD.[10] A much smaller number of investigations of social cognition[11–16] appear generally suggestive of the same, but small sample sizes, non-tract-specific white matter estimates,[17] and narrow social cognitive assessments[18] are limitations. Moreover, all social cognitive studies implemented a case-control design and univariate statistics, which may together preclude biomarker identification.[19]

The objective of the present study was to illuminate the relationship between neurocognitive and social cognitive performance, and estimates of white matter integrity, in individuals with an SSD and healthy control comparisons (HC). We asked: are there some white matter tracts that relate only to neurocognition, others only to social cognition, and still others to both? To answer this question, we harnessed data from the Research Domain Criteria (RDoC) study "Social Processes Initiative in the Neurobiology of the Schizophrenia(s)" (SPINS), which collected data from 2015 to 2019. SPINS was designed to identify impairments in neural circuit structure giving rise to social cognitive deficits, across the continuum of people with and without an SSD. An advantage of the SPINS study is its broad testing of social cognition. Administering several independent tasks was necessary because, at the initiation of the SPINS study and still today, no standardized battery of social cognition exists, akin to the MATRICS (Measurement and Treatment Research to Improve Cognition in Schizophrenia) Consensus Cognitive Battery (MCCB) for neurocognition.[20] Though other studies from our group have used the SPINS dataset to investigate the factor structure of social cognition,[21] and illuminate the relationship between cognitive performance and functional activity and connectivity[7,22,23] the present investigation evaluates associations among white matter, neurocognition, and social cognition.

We employed multivariate canonical correlation analysis (CCA)[24,25] to reveal the joint structure of white matter fractional anisotropy estimates, and social cognitive and neurocognitive performance scores. CCA was an attractive method by which to approach our question, because it allows for the identification of cognitively-relevant brain features beyond "one-tract-one-function" associations, and also lends itself easily to dimensional analysis (cf. case-control analysis), in keeping with the RDoC paradigm.[26] Via exploratory analyses, we explored the utility of CCA model derivatives to illuminate three topical debates in the field; namely if cognition-constrained estimates of white matter microstructure might reveal novel subgroups, confirm psychiatric diagnosis, and/or predict cross-sectional functional outcomes.

Methods

Participants

The SPINS study recruited participants across three centers: the Center for Addiction and Mental Health (Toronto), the Maryland Psychiatric Research Center (Maryland), and Zucker Hillside Hospital (New York). To allow for non-exact in-sample replication, we derived a "discovery" and "validation" sample in accordance with the MRI model upon which participants' imaging data were acquired. Our discovery sample comprised $n = 135$ (85 SSD) collected on a single 3T GE 750w Discovery in Toronto, and our validation sample comprised $n = 173$ (93 SSD) collected at all three centers on prospectively harmonized 3 T Siemens Prismas.

Eligibility Criteria. Participants with SSDs met DSM-5 diagnostic criteria for schizophrenia, schizoaffective disorder, schizophreniform disorder, delusional disorder, or unspecified schizophrenia spectrum and other psychotic disorder, assessed using an adapted Structured Clinical Interview for DSM (SCID-IV-TR). Individuals with SSD were symptomatically stable and had no change in antipsychotic medication or decrement in functioning in 4 weeks before enrollment. Exclusion criteria included a history of head trauma resulting in unconsciousness, intellectual disability, substance use disorder within the past 3 months, debilitating or unstable medical illness, neurological disease, and MR contraindications. Additionally, HCs did not ever have a DSM-IV Axis I disorder, excepting adjustment disorder, phobic disorder, past major depressive disorder (over 2 years prior; presently unmedicated), or a first-degree relative with a primary psychotic disorder. Participants ranged in age between 18 and 59. The protocol was approved by the respective research ethics boards and institutional review boards, and all participants provided written informed consent. All research was conducted in accordance with the Declaration of Helsinki.

Participant Assessment

Neurocognitive Measures. Neurocognition was evaluated using the MATRICS MCCB,[20] which provides domain scores for processing speed, reasoning and problem-solving, attention/vigilance, working memory, and verbal and visual learning. We omitted the social cognition domain, given our social cognitive battery.[27]

Social Cognitive Measures. We tested social cognition via five tasks: the Penn Emotion Recognition Test (ER-40) which tests facial emotion recognition[28]; the Reading the Mind in the Eyes Task (RMET) which tests mental state inference from the eyes[29]; the Empathic Accuracy (EA) task which involves evaluation of positive or negative emotions via video vignettes[30,31]; the Relationships Across Domains (RAD) task which requires understanding of interpersonal relations[32] and The Awareness of Social Inference Test-Revised (TASIT) which assesses emotion and social inference via video vignettes.[33] Validation studies have found these tasks to be fit for clinical trial use,[30,34] with the exception of the RAD, which is a test of social perception with adequate psychometric properties.

Other Measures. Psychiatric symptoms were evaluated in the SSD sample using the Brief Psychiatric Rating Scale (BPRS)[35] and the Scale for the Assessment of Negative Symptoms (SANS).[36] In both SSD and HC groups, the Birchwood Social Functioning Scale (BSFS)[37] evaluated social functioning, and the Cumulative Illness Rating Scale-Geriatric (CIRS-G)[38] evaluated chronic illness burden. In the SSD group only, we assessed functioning via the Quality of Life Scale (QLS),[39] extra-pyramidal signs via the Simpson-Angus Scale (SAS),[40] and chlorpromazine equivalents (CPZE)[41] for antipsychotics.

Imaging Procedures

Diffusion Imaging: Acquisition and Preprocessing. We acquired a high-angular axial EPI dual spin echo sequence diffusion scan.[42] Parameters were prospectively harmonized across scanners within the limits of hardware, as follows: 60 gradient directions, $b = 1000$, 5 $b = 0$ images (two scanners 6 $b = 0$ s), TR = 8800 ms (one scanner TR = 17700 ms), TE = 85 ms, FOV = 256 mm; in-plane matrix 128 × 128, and 2.0 mm isotropic voxels. All images were pre-processed identically: (1) brain masking via two-step agreement in AFNI (BET) and MRtrix3 (dwi2mask), (2) motion correction for inter- and intra-volume movement via FSL (eddy), and (3) susceptibility distortion correction via BrainSuite (BDP).

White Matter Analysis. We fit a tensor and reconstructed white matter tracts via deterministic unscented Kalman filter tractography,[43] using the "WhiteMatterAnalysis" algorithm available in 3D Slicer (https://github.com/SlicerDMRI). We clustered fibers via supervised groupwise registration[44] to the ORG (O'Donnell Research Group) atlas.[45,46] We report fractional anisotropy (FA), as it is the most commonly-reported diffusion index,[47] and reflects the most disruption in both illness and cognitive impairment.[48] We confirmed that no scanner effect was evident in FA values across the three scanners in the validation sample. For imaging quality control procedures, see Supplementary material S2.

Statistical Procedures

Preprocessing. For all variables of primary interest (ie, estimates of neurocognitive and social cognitive performance, and white matter microstructure), we (1) removed outliers via the adjusted boxplot method (< 3% of values),[49] (2) imputed removed and missing data (< 0.05% of values) with chained equations,[50] (3) transformed skewed distributions via the Yeo-Johnson power transformation (all social cognition variables, all negatively skewed),[51] and (4) ensured that no variables were multicollinear (all in-set VIF < 6).[52] Lastly, we residualized non-meaningful sources of variation on white matter microstructure; namely age, sex, and antipsychotic medication load as estimated by CPZE.[53–55]

Canonical Correlation Analysis. We employed CCA[24] to model the "doubly multivariate" associations of white matter microstructure (X set), and neurocognition and social cognition (Y set). Our X set comprised FA estimates in 19 deep white matter tracts, selected on the basis of a previously demonstrated connection to neurocognition and/or social cognition in existing literature, and reliable tract segmentation (Supplementary material S1). Our Y set comprised the previously described six MATRICS MCCB domain scores, and 10 social cognition scores: total scores from the ER-40, RMET, EA, and RAD, and subscale scores from the TASIT (TASIT 1; TASIT 2: sincere; TASIT 2: paradoxical sarcasm; TASIT 2: simple sarcasm; TASIT 3: lies; TASIT 3: sarcasm). Though 6 of our social cognition scores derive from the same test, we have previously shown each to capture unique variance.[21] Thus, our total of 35 features across sets meets the recommended 5:1 observation-to-feature ratio in our $n = 173$ replication sample.[56] It is not considered inherently problematic that our Y set contains features of different types (namely domain scores, total scores, and subscale scores).[25]

CCA employs an unsupervised matrix decomposition technique to re-express X and Y features as lower-dimensional "canonical variates", X' and Y', that are maximally correlated under the constraint of orthogonality. CCA uses a nested procedure to test for significant associations between variates, and as such, it does not require correction for multiple comparisons.[57] CCA's primary outcome metric is a canonical correlation value (Rc), which estimates shared structure across variates. Of interest to us, CCA provides interpretable estimates of feature importance via structure coefficients (r_s), which express the univariate correlation between a given feature and its canonical variate. Because estimates can prove unstable across samples,[58–60] we interpreted features surpassing the conservative threshold of $|r_s|_{\geq 0}$.45 in both the discovery and validation samples.[61]

Exploratory Analyses. An additional output of CCA are participant-wise "variate scores", for each variate, and each set, which weight observed values by the model's coefficients. Thus, variate scores have the interesting property that they are constrained to lie along axes of variance maximally related to the other set.[62] In our case, X' variate scores capture participant-wise FA estimates adjusted for cognitive performance, which we refer to subsequently as "cognition-constrained white matter". In three exploratory analyses, we probed the utility of cognition-constrained white matter scores from significant variates to (1) illuminate natural subgroups (via clustering), (2) confirm "ground truth" diagnostic labels (via classification), and (3) predict social functioning (via regression). See Supplementary materail S3 for a complete description of exploratory methods.

Results

Participant Characteristics

Table 1 summarizes participant demographic, clinical, neurocognitive, and social cognitive characteristics. FA values in the discovery and validation samples are available in Supplementary material S4. We observed small negative to large positive bivariate correlations within and between white matter and cognition estimates, shown in Supplementary material S5.

CCA Analyses: White Matter–Cognition Relationships

The CCA analyses were conducted identically in the discovery and validation samples. The full models showed high canonical correlation values [$Rc_{DISCOVERY} = 0.71$; $Rc_{VALIDATION} = 0.72$] (figure 1A). Permutation against empirical null distributions found both models to be significant: [$p_{DISCOVERY} < 0.005$; $p_{VALIDATION} \leq 0.001$] (figure 1B), as was parametric testing via the asymptotic Hotelling-Lawley Trace statistic [$p_{DISCOVERY} = 0.027$; $p_{VALIDATION} \leq 0.001$], though the more commonly Wilks' lambda statistic was mixed [$p_{DISCOVERY} = 0.065$, $p_{VALIDATION} \leq 0.001$]. In both samples, nested hierarchical significance testing revealed only the first canonical variate pair (CV1) of 16 to be significant. CV1 explained a substantial portion of variance ($Rc^2_{DISCOVERY} = 50\%$, $Rc^2_{VALIDATION} = 23\%$) and redundancy ($Rd_{DISCOVERY} = 52\%$, $Rd_{VALIDATION} = 22\%$). Sensitivity analyses in which we systematically altered aspects of our statistical preprocessing regime (ie, outlier removal, imputation, and/or normality correction) did not change the nature of the global CCA results, and jacknife resampling (ie, iterative participant removal) showed global results to be stable. Residual analysis of participants' scores on CV1 showed an indistinguishable pattern across SSD and HC.

CV1 showed several features bearing canonical loadings beyond our chosen "importance" threshold of $|r_s| \geq 0.45$. Comparison of important features in the discovery and validation samples showed similarity in polarity and magnitude, though differences were evident in precise value and rank (figure 1C). In both samples, the body of the corpus callosum (CC3) and the right uncinate fasciculus (UF) contributed highly, as did the MCCB speed of processing, attention and vigilance, verbal learning, and visual learning (neurocognition), and TASIT 3 sarcasm (social cognition).

Exploratory Analyses

Next, we conducted three exploratory analyses using participant-wise cognition-constrained white matter scores from CV1. Because these values are standardized, we combined them across the discovery and validation samples, and derived a training ($n = 200$) and testing ($n = 108$) sample.

Cognition-constrained White Matter and Clustering. First, we evaluated the potential of cognition-constrained white matter to reveal participant subgroups, by clustering the training set via Ward's complete-linkage method. We found a five-cluster solution was optimal (figure 2A), with a Calinski-Harabasz index (CHI) of 595.71 (figure 2B). However, permutation testing showed this CHI was likely to occur under the null hypothesis of no clusters embedded in the data ($P = .105$) (figure 2C).[62,63]

Cognition-Constrained White Matter and Agreement With Case/Control Designation. Second, we applied receiver operating characteristic (ROC) curve analysis to determine if cognition-constrained white matter might reveal a clear diagnostic cut-point, separating SSD and HC. The ROC curve in the training set showed excellent recovery (AUC = 0.941 [0.917–0.965]) (figure 2D), which was highly unlikely to arise from chance ($D = 10.36$, $P < .001$). The Youden index identified an optimal cut-point at $X'_1 = -0.237$ (AUC = 0.939 [0.908–0.970], balanced accuracy = 88.5%), which showed excellent predictive ability when applied to the held-out test sample (AUC = 0.948 [0.903–0.982], balanced accuracy = 87%) (figure 2E). Misclassifications of participants with SSD and HCs were equally likely (McNemar's test), and misclassified participants were not differentiated by age or sex, nor symptom severity (SSD only). Comparison of this model to alternatives taking each white matter feature uninfluenced by cognition as a predictor, as well as their combined one-dimensional representation (PCA), found that only the cognition-constrained classifier achieved "exceptional" performance.

Cognition-Constrained White Matter and Prediction of Real-world Functioning. Lastly, we evaluated if cognition-constrained white matter might predict social functioning, as measured by the BSFS,[37] using 5-fold

Table 1. Participant Demographic, Clinical, Neurocognitive, and Social Cognitive Characteristics

	Discovery sample (GE Discovery)				Validation sample (Siemens Prismas)				Sample comparison
	SSD-HC	SSD	HC		SSD-HC	SSD	HC		
	n = 135	*n* = 89	*n* = 46	*P*-value adj.[†]	*n* = 173	*n* = 91	*n* = 82	*P*-value adj.[†]	*P*-value adj.[‡]
	Mean (SD)	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	Mean (SD)		
Demographics									
Sex (F:M)	48:87	28:61	20:26	–	70:103	29:62	41:41	–	–
Age	27.51 (7.68)	27.69 (7.36)	27.17 (8.34)	1.000	32.65 (10.56)	32.27 (10.67)	33.06 (10.48)	1.000	.005
Handedness	0.67 (0.45)	0.65 (0.47)	0.70 (0.41)	1.000	0.67 (0.44)	0.72 (0.38)	0.61 (0.50)	.661	1.000
Education	14.19 (2.13)	13.52 (1.97)	15.48 (1.82)	.000	14.99 (2.38)	13.96 (2.14)	16.13 (2.10)	.000	.024
Parental education	15.73 (2.96)	15.44 (3.15)	16.26 (2.49)	.728	15.42 (3.07)	15.17 (3.34)	15.69 (2.76)	1.000	1.000
WTAR	113.04 (11.78)	111.52 (12.73)	116.00 (9.07)	.139	109.28 (12.28)	105.32 (12.10)	113.48 (11.11)	.004	.216
Clinical									
BPRS	–	29.96 (6.65)	–	–	–	31.16 (7.88)	–	–	1.000
SANS	–	2.57 (2.23)	–	–	–	2.32 (2.63)	–	–	.754
BSFS	151.09 (27.02)	140.71 (23.66)	171.17 (21.29)	.000	154.06 (29.94)	134.47 (23.80)	175.79 (19.10)	.000	1.000
QLS	–	79.77 (20.03)	–	–	–	69.03 (20.57)	–	–	.006
CIRS-G	2.54 (2.27)	2.87 (2.57)	1.91 (1.35)	.038	2.89 (3.09)	4.10 (3.42)	1.55 (1.97)	.000	1.000
SAS	–	22.58 (12.09)	–	–	–	25.96 (12.07)	–	–	1.000
CPZE	–	321.96 (290.86)	–	–	–	523.11 (444.68)	–	–	.005
Neurocognition									
Speed of processing	45.38 (12.52)	41.84 (12.31)	52.22 (9.87)	.000	45.86 (14.52)	38.30 (13.75)	54.26 (10.09)	.000	1.000
Attention/vigilance	41.41 (11.26)	38.10 (10.80)	47.80 (9.29)	.000	43.88 (14.22)	40.58 (12.36)	47.55 (15.29)	.021	1.000
Working memory	45.22 (11.97)	41.63 (11.25)	52.17 (10.21)	.000	44.65 (11.52)	42.57 (10.64)	46.95 (12.08)	.203	1.000
Verbal learning	44.35 (9.58)	41.87 (9.06)	49.15 (8.78)	.000	45.14 (11.13)	39.93 (9.11)	50.93 (10.33)	.000	1.000
Visual learning	45.01 (11.28)	42.42 (11.97)	50.04 (7.70)	.000	42.36 (12.84)	37.26 (12.26)	48.02 (11.02)	.000	.882
Reasoning and problem-solving	44.64 (9.61)	43.17 (9.71)	47.48 (8.85)	.174	46.30 (11.61)	43.35 (12.19)	49.57 (10.04)	.005	1.000
Social cognition									
EA	0.84 (0.23)	0.79 (0.26)	0.93 (0.13)	.002	0.79 (0.26)	0.63 (0.30)	0.83 (0.26)	.000	.005
ER-40	−2038.66 (573.77)	−2184.42 (632.07)	−1756.65 (275.21)	.000	−2184.42 (632.07)	−2342.35 (681.08)	−1848.78 (374.13)	.000	1.000
RAD	56.27 (9.43)	53.69 (9.91)	61.26 (5.85)	.000	53.69 (9.91)	51.91 (8.68)	60.18 (5.61)	.000	1.000
RMET	26.51 (4.48)	25.57 (4.71)	28.33 (3.35)	.002	25.57 (4.71)	24.63 (5.32)	27.13 (4.07)	.009	1.000
TASIT 1	24.19 (2.78)	23.91 (3.06)	24.72 (2.06)	1.000	23.91 (3.06)	22.16 (3.31)	24.79 (2.25)	.000	.362
TASIT 2 paradoxical sarcasm	17.22 (3.30)	16.47 (3.56)	18.67 (2.07)	.000	16.47 (3.56)	15.40 (3.99)	18.29 (2.26)	.000	1.000
TASIT 2 simple sarcasm	16.28 (4.13)	15.33 (4.60)	18.13 (2.04)	.000	15.33 (4.60)	15.22 (4.63)	18.62 (1.97)	.000	1.000
TASIT 2 sincere	17.18 (2.81)	17.10 (2.73)	17.33 (2.99)	1.000	17.10 (2.73)	16.81 (3.46)	17.49 (2.53)	1.000	1.000

Table 1. Continued

	Discovery sample (GE Discovery)				Validation sample (Siemens Prismas)				Sample comparison
	SSD-HC	SSD	HC		SSD-HC	SSD	HC		
	n = 135	*n* = 89	*n* = 46	*P*-value adj.†	*n* = 173	*n* = 91	*n* = 82	*P*-value adj.†	*P*-value adj.‡
	Mean (SD)	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	Mean (SD)		
TASIT 3 lies	26.12 (4.20)	25.03 (4.44)	28.22 (2.67)	.000	25.03 (4.44)	24.53 (4.94)	26.38 (4.14)	.130	1.000
TASIT 3 sarcasm	25.47 (4.58)	24.52 (4.66)	27.33 (3.82)	.005	24.52 (4.66)	23.32 (5.54)	27.63 (3.10)	.000	1.000

Note: All *P*-values for continuous variables are derived from a t-test, and the categorical variable (sex) from a chi-square test. The three "*P*-value adj" columns have been Bonferroni corrected for multiple comparisons. ***Units***: Age, education, and parental education are reported in years; Handedness 0 = left, 1 = right; CPZE is mg/day. The SANS score excludes the Attention subscale, and the BSFS excludes the Employment score. All neurocognitive scores are total scores on the MATRICS MCCB domains. EA is accuracy. ER-40 is test completion time in seconds (introverted scores, so higher is better). RAD, RMET, and TASIT are total correct.
BPRS, Brief Psychiatric Rating Scale; BSFS, Birchwood Social Functioning Scale; CIRS-G, Cumulative Illness Rating Scale-Geriatric; CPZE, chlorpromazine equivalents; EA, Empathic Accuracy; ER-40, Penn Emotion Recognition test; F, female; HC, healthy control; M, male; QLS, Quality of Life Scale; RAD, Relationships Across Domains; SANS, Scale for the Assessment of Negative Symptoms; SAS, Simpson-Angus Scale; TASIT, The Awareness of Social Inference Test-Revised.
†*P*-values within the Discovery sample and Validation sample "blocks" respectively compare the SSD and HC scores (providing a comparison of participant groups within sample), and the respective SSD-HC columns describe the sample's combined participants' mean and standard deviation.
‡*P*-values in the Sample comparison block compare the combined SSD-HC means and standard deviations between the Discovery and Validation samples (providing a comparison across samples).

5-repeat cross-validated linear regression. In the test set, cognition-constrained white matter explained 25.3% in BSFS variation (adjusted $R^2 = 0.253$), with prediction error of RMSE = 25.082 in BSFS units (maxmin normalized RMSE = 0.169) (figure 2F). A non-nested likelihood ratio test comparing this model to an alternative leveraging all white matter features uninfluenced by cognition as predictors found the models were distinguishable ($P = .004$), and that the alternative model uninfluenced by cognition possessed better goodness-of-fit ($P < .001$).

Discussion

Understanding structure-cognition relationships is an important focus of schizophrenia research and cognitive neuroscience more broadly. Here, we sought to understand if the microstructural integrity of shared (or distinct) white matter tracts is associated with neurocognitive and social cognitive performance. Multivariate CCA of FA estimates in 19 tracts (*X* set), and 6 neurocognitive and 10 social cognitive performance scores (*Y* set) revealed a significant structure-cognition association, stable in the face of participant- and feature-wise perturbation. Subsequent examination of model coefficients (r_s values) revealed that neurocognition (MCCB processing speed, attention and vigilance, verbal learning, and visual learning) and social cognition (TASIT 3 sarcasm) were subserved by common tracts, namely the body of corpus callosum and right UF. We did not observe evidence of distinct circuitry relating to neurocognition vs social cognition separately. The finding that neurocognitive and social cognitive performance relies on partially overlapping circuitry complements functional neuroimaging evidence of common cognitive processing strategies in healthy individuals[64–67] and the SPINS sample.[7,22,23]

Though the CCA illuminated shared cognitive circuitry, it is notable that more neurocognitive (4) than social cognitive (1) features shared this mapping. This may suggest that neurocognition is more broadly reliant than social cognition on white matter integrity. However, it may also reflect that the employed social cognitive tasks lack sensitivity to neurobiology; after all, none were devised with consideration of how (or if) they might map to brain circuits. It is plausible that different social cognitive tasks, or broader dimensions/domains, might prove more proximal to white matter abnormalities. An alternative explanation is that the TASIT 3 sarcasm may be "more neurocognitive" than other social cognition tasks.[68] Resolving questions such as these will be greatly aided by a consensus social cognitive battery, akin to the MCCB.[69]

The CCA model highlighted high contributions from the body of corpus callosum and right UF, providing motor interhemispheric and orbito-frontal/medial prefrontal-amygdalar connections, respectively. Alterations to the corpus callosum are among the most robust findings in

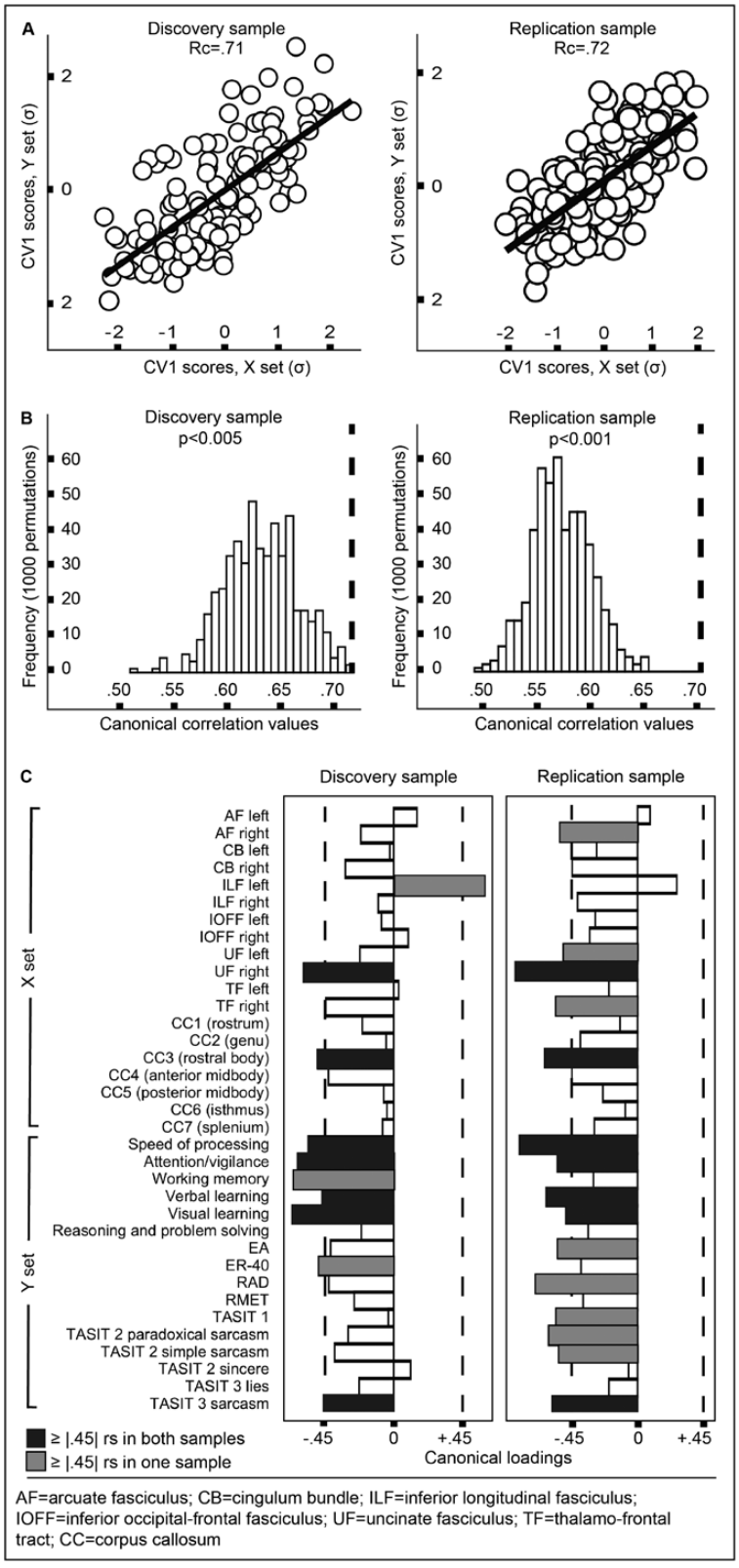


Fig. 1. White matter–cognition relationships. (A) Full model canonical correlations (Rc) and participant-wise scores on the first canonical variate (CV1), in the discovery (left) and replication (right) samples. (B) The observed canonical correlation values (dashed lines) are unlikely to arise from chance. (C) Several *X* and *Y* set variables were important to the CCA resolution as established by the conversative threshold of $|r_s|_{\geq} 0.45$ (dashed lines). Estimates were grossly similar in magnitude and polarity across samples. *Note*: Rc, canonical correlation; CV1, first canonical variate; rs, standardized structure coefficient (canonical loading. For *X* and *Y* set abbreviations in (C), consult paper text.

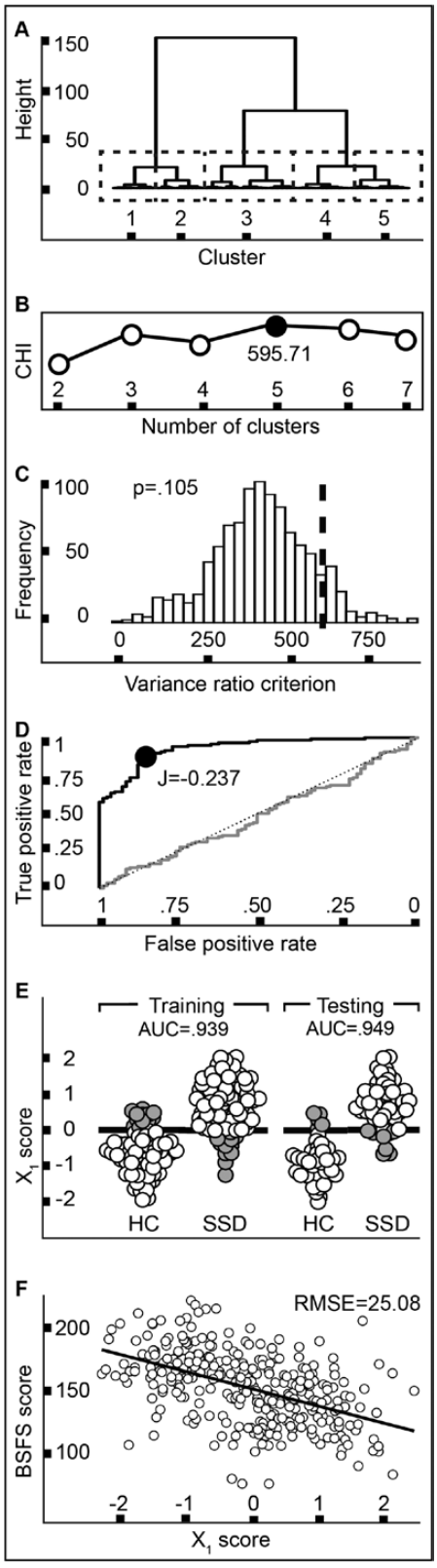


Fig. 2. Exploratory results. (A–C) clustering analysis; (D, E) classification analysis, (**F**) prediction analysis. (A) The dendrogram in the training sample. The five-cluster solution is highlighted (dashed boxes). (B) The five-cluster solution demonstrated the highest Calinski-Harabasz index (black dot). (C) Permutation testing showed the observed CHI (dashed line) was not distinguishable from a null distribution (D). A Receiver Operating Characteristic curve analysis in the training sample showed excellent recovery, and the Youden index (*J*) identified an optimal cut-point (black dot). (E) In both the training and testing samples, the optimal cut-point showed excellent predictive ability, in both diagnostic groups (correct classifications in white, misclassifications in gray). (F) X'_1 scores in the training set were predictive of social functioning. *Note:* CHI, Calinski-Harabasz index; J, Youden index; X'_1, participant-wise cognition-constrained white matter scores on the first canonical variate; AUC, area-under-the-curve; HC, healthy control; SSD, schizophrenia spectrum disorder; BSFS, Birchwood Social Functioning Scale; RMSE, root-mean-square-error.

schizophrenia[70] and have been related to the cognitions. For instance, one recent study found that integrity of the body of the corpus showed the strongest association to working memory and processing speed (the only two neurocognitive variables analyzed),[71] and another found that integrity of the body of the corpus was most related to social cognition, albeit only assessed via one task believed to measure social perception.[72] Our results corroborate and extend these findings by demonstrating the role of the body of the corpus in the same participants, across a diverse battery of cognitive tasks.

The UF is also known to be microstructurally disturbed in schizophrenia,[73] but only a small body of empirical work has explored its relation to the cognitions. One study found that integrity of the bilateral uncinate was positively correlated with several neurocognitive domains, as well as emotion processing.[74] Another study found that integrity of the right uncinate was associated with social perception, but the correlation was negative.[75] Our results underscore that the uncinate is important to the cognitions, though further research is needed to understand an apparent disagreement in lateralization,[76] as well as possible specialization of subcomponents.[77]

Latent structure-cognition associations were statistically indistinguishable between SSD and HC participants. However, this does not prove that there are no unique structural marker(s) of impaired cognition; we did not test for this, and others have found some evidence in favor of it in schizophrenia, for example,[71] This result extends prior work by our group, which demonstrated that the statistical structure of cognition alone is invariant in SSD and HC.[21] Indistinguishable structure-cognition associations provide post hoc endorsement of a dimensional relationship across SSD and HC, and extends to structure prior evidence from the SPINS study that multivariate function-cognition links may be better described as deficit-specific, as opposed to diagnosis-specific.[7,22,23]

In exploratory analyses, we leveraged CCA model outputs (ie, X'_1 scores representing "cognition-constrained white matter") to perform clustering, classification, and prediction analyses. Our clustering analysis failed to illuminate natural subgroups, that is, "biotypes". This may

be the consequence of our a priori decision to limit our search within scores derived from significant canonical variate(s): we thus searched within a one-dimensional subspace which is likely to capture something of a global brain-behavior relationship. Our classification analysis found that cognition-constrained white matter served as an excellent diagnostic biomarker (balanced accuracy in test set = 87%), with confusion mostly circumscribed to a fuzzy "tipping point".[78] This performance was favorable among the 10 prior studies that have sought to distinguish individuals with schizophrenia from HCs on the basis of white matter features alone, with accuracies reported from 62 to 94%.[79–88] Finally, our regression analysis showed that cognition-constrained white matter predicted cross-sectional social functioning scores, though an alternative model utilizing white matter alone showed superior performance. This may reflect the fact that cognition-constrained white matter is constrained by neurocognitive performance, which we have previously found to be less related to social functioning than social cognitive performance.[21] Despite these mixed exploratory results, we view weighted structure-cognition scores to be of high utility to various clustering, classification, and prediction applications. In particular, our finding that cognition-constrained white matter accurately predicts diagnosis could be useful in advancing efforts into the prodrome or prior, given that subtle differences in both white matter[89] and cognitive performance[90,91] are evident before frank psychosis onset.

Our study has several limitations. Pertaining to the SPINS sample, participants were heterogeneous across many domains, including those that may influence neurocognition, social cognition, and white matter microstructure, including age,[92,93] duration of illness,[94–96] and antipsychotic exposure.[97,98] Further, participants with SSD were disproportionately male.[99,100] We attempted to mitigate these limitations by adjusting primary outcomes for age, sex, and CPZE.A second set of limitations pertain to our use of CCA. CCA is "data hungry" in that it requires a high observation-to-feature ratio to avoid overfitting (ie, identifying spurious associations that fail to generalize). To achieve an adequate observation-to-feature ratio of approximately 5:1[56] in our validation sample, we opted to limit our X set to 19 tracts. This feature selection undercuts the full data-driven power of CCA. It is possible that other tracts, perhaps the cerebellar (peduncles) and projection tracts, may prove relevant to cognition in schizophrenia, based on analogous findings in healthy individuals.[101,102] A related limitation is that estimates of feature importance may be especially unstable across samples,[58–60] with a recent suggestion that this instability only resolves with an observation-to-feature ratio of 50:1,[103] which is 10-fold that of our and most other imaging-cognition studies. We attempted to mitigate this worry by imposing a high threshold for interpretation (canonical loadings $|r_s|_{\geq 0}.45$),[61] and interpreting only those features surpassing this high threshold in our two samples.

Caveats notwithstanding, this study confirms that white matter microstructure captures an important latent component of neurocognitive and social cognitive performance, and provides novel evidence that neurocognitive and social cognitive performance are subserved by common white matter circuitry. Our results are strengthened by our comprehensive cognitive batteries, use of a multivariate approach, and in-sample replication. Future work should probe the effect of targeting the body of the corpus and the right UF: improved microstructural integrity might bring about enhanced cognitive ability and a corresponding improvement to functional outcomes in SSD.

Supplementary Material

Supplementary material is available at https://academic.oup.com/schizophreniabulletin/.

Funding

This work was supported by the National Institute of Mental Health (grant numbers 1/3R01MH102324–01 to ANV, 2/3R01MH102313–01 to AKM, and 3/3R01MH 102318–01 to RWB).

References

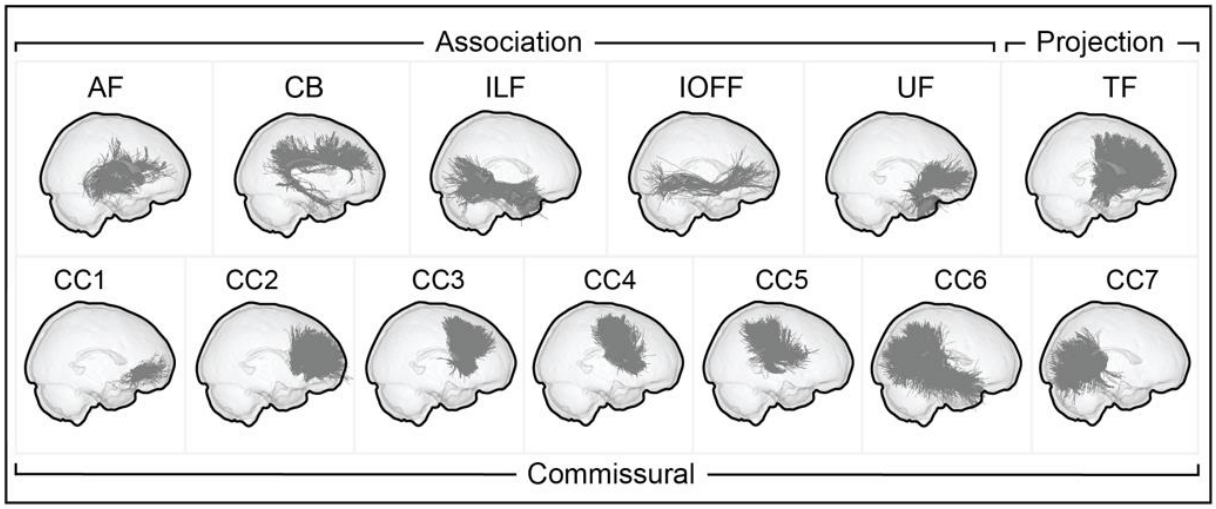
1. McCleery A, Nuechterlein KH. Cognitive impairment in psychotic illness: prevalence, profile of impairment, developmental course, and treatment considerations. *Dialogues Clin Neurosci.* 2019;21:239–248.
2. Vaskinn A, Horan WP. Social cognition and schizophrenia: unresolved issues and new challenges in a maturing field of research. *Schizophr Bull.* 2020;46:464–470.
3. Green MF, Horan WP, Lee J. Nonsocial and social cognition in schizophrenia: current evidence and future directions. *World Psychiatry.* 2019;18:146–161.
4. Fett A-KJ, Viechtbauer W, Dominguez M-G, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. *Neurosci Biobehav Rev.* 2011;35:573–588.
5. Halverson TF, Orleans-Pobee M, Merritt C, Sheeran P, Fett A-K, Penn DL. Pathways to functional outcomes in schizophrenia spectrum disorders: meta-analysis of social cognitive and neurocognitive predictors. *Neurosci Biobehav Rev.* 2019;105:212–219.
6. Sergi MJ, Rassovsky Y, Widmark C, *et al*. Social cognition in schizophrenia: relationships with neurocognition and negative symptoms. *Schizophr Res.* 2007;90:316–324.
7. Viviano J, Buchanan RW, Calarco N, *et al*. Resting-state connectivity biomarkers of cognitive performance and social function in individuals with schizophrenia spectrum disorder and healthy control subjects. *Biol Psychiatry.* 2018;84:665–674.

8. Deckler E, Hodgins GE, Pinkham AE, Penn DL, Harvey PD. Social cognition and neurocognition in schizophrenia and healthy controls: intercorrelations of performance and effects of manipulations aimed at increasing task difficulty. *Front Psychiatry.* 2018;9:356.
9. Wu W, Miller KL. Image formation in diffusion MRI: a review of recent technical developments. *J Magn Reson Imaging.* 2017;46:646–662.
10. Holleran L, Kelly S, Alloza C, *et al*. The relationship between white matter microstructure and general cognitive ability in patients with schizophrenia and healthy participants in the ENIGMA consortium. *Am J Psychiatry.* 2020;177:537–547.
11. Fujiwara H, Hirao K, Namiki C, *et al*. Anterior cingulate pathology and social cognition in schizophrenia: a study of gray matter, white matter, and sulcal morphometry. *Neuroimage.* 2007;36:1236–1245.
12. Miyata J, Yamada M, Namiki C, *et al*. Reduced white matter integrity as a neural correlate of social cognition deficits in schizophrenia. *Schizophr Res.* 2010;119:232–239.
13. Fujino J, Takahashi H, Miyata J, *et al*. Impaired empathic abilities and reduced white matter integrity in schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry.* 2014;48:117–123.
14. Zhao X, Sui Y, Yao J, *et al*. Reduced white matter integrity and facial emotion perception in never-medicated patients with first-episode schizophrenia: a diffusion tensor imaging study. *Prog Neuropsychopharmacol Biol Psychiatry.* 2017;77:57–64.
15. Saito Y, Kubicki M, Koerte I, *et al*. Impaired white matter connectivity between regions containing mirror neurons, and relationship to negative symptoms and social cognition, in patients with first-episode schizophrenia. *Brain Imaging Behav.* 2018;12:229–237.
16. Kim NS, Lee TY, Hwang WJ, *et al*. White matter correlates of theory of mind in patients with first-episode psychosis. *Front Psychiatry.* 2021;12:617683–617683.
17. Bach M, Laun FB, Leemans A, *et al*. Methodological considerations on tract-based spatial statistics (TBSS). *Neuroimage.* 2014;100:358–369.
18. Pinkham AE, Penn DL, Green MF, Harvey PD. Social cognition psychometric evaluation: results of the initial psychometric study. *Schizophr Bull.* 2016;42:494–504.
19. McGuire P, Sato JR, Mechelli A, Jackowski A, Bressan RA, Zugman A. Can neuroimaging be used to predict the onset of psychosis? *Lancet Psychiatry.* 2015;2:1117–1122.
20. Nuechterlein KH, Green MF, Kern RS, *et al*. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J Psychiatry.* 2008;165:203–213.
21. Oliver LD, Haltigan JD, Gold JM, *et al*. Lower- and higher-level social cognitive factors across individuals with schizophrenia spectrum disorders and healthy controls: relationship with neurocognition and functional outcome. *Schizophr Bull.* 2018. doi:10.1093/schbul/sby114.
22. Hawco C, Buchanan RW, Calarco N, *et al*. Separable and replicable neural strategies during social brain function in people with and without severe mental illness. *Am J Psychiatry.* 2019. doi:10.1176/appi.ajp.2018.17091020.
23. Oliver LD, Hawco C, Homan P, *et al*. Social cognitive networks and social cognitive performance across individuals with schizophrenia spectrum disorders and healthy control participants. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2020. doi: 10.1016/j.bpsc.2020.11.014.
24. Hotelling H. Relations between two sets of variates. *Biometrika.* 1936;28:321.
25. Wang H-T, Smallwood J, Mourao-Miranda J, *et al*. Finding the needle in a high-dimensional haystack: canonical correlation analysis for neuroscientists. *Neuroimage.* 2020;216:116745.
26. Cuthbert BN, Insel TR. Toward new approaches to psychotic disorders: the NIMH Research Domain Criteria project. *Schizophr Bull.* 2010;36:1061–1062.
27. Green MF, Harris JG, Nuechterlein KH. The MATRICS consensus cognitive battery: what we know 6 years later. *Am J Psychiatry.* 2014;171:1151–1154.
28. Kohler CG, Bilker W, Hagendoorn M, Gur RE, Gur RC. Emotion recognition deficit in schizophrenia: association with symptomatology and cognition. *Biol Psychiatry.* 2000;48:127–136.
29. Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I. The "Reading the Mind in the Eyes" test revised version: a study with normal adults, and adults with asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry.* 2001;42:241–251.
30. Kern RS, Penn DL, Lee J, *et al*. Adapting social neuroscience measures for schizophrenia clinical trials, part 2: trolling the depths of psychometric properties. *Schizophr Bull.* 2013;39:1201–1210.
31. Olbert CM, Penn DL, Kern RS, *et al*. Adapting social neuroscience measures for schizophrenia clinical trials, part 3: fathoming external validity. *Schizophr Bull.* 2013;39:1211–1218.
32. Sergi MJ, Fiske AP, Horan WP, *et al*. Development of a measure of relationship perception in schizophrenia. *Psychiatry Res.* 2009;166:54–62.
33. McDonald S, Flanagan S, Rollins J. *The Awareness of Social Inference Test (Revised)*. Sydney, Australia: Pearson Assessment; 2011.
34. Pinkham AE, Harvey PD, Penn DL. Social cognition psychometric evaluation: results of the final validation study. *Schizophr Bull.* 2018;44:737–748.
35. Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychol Rep.* 1962;10:799–812.
36. Andreasen NC. Negative symptoms in schizophrenia. Definition and reliability. *Arch Gen Psychiatry.* 1982;39:784–788.
37. Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S. The Social Functioning Scale. The development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. *Br J Psychiatry.* 1990;157:853–859.
38. Miller MD, Paradis CF, Houck PR, *et al*. Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. *Psychiatry Res.* 1992;41:237–248.
39. Heinrichs DW, Hanlon TE, Carpenter WT Jr. The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. *Schizophr Bull.* 1984;10:388–398.
40. Janno S, Holi MM, Tuisku K, Wahlbeck K. Validity of Simpson-Angus Scale (SAS) in a naturalistic schizophrenia population. *BMC Neurol.* 2005;5:5.
41. Leucht S, Samara M, Heres S, Davis JM. Dose equivalents for antipsychotic drugs: the DDD method. *Schizophr Bull.* 2016;42(Suppl 1):S90–S94.
42. Tuch DS, Reese TG, Wiegell MR, Makris N, Belliveau JW, Wedeen VJ. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn Reson Med.* 2002;48:577–582.

43. Malcolm JG, Shenton ME, Rathi Y. Filtered multitensor tractography. *IEEE Trans Med Imaging.* 2010;29:1664–1675.
44. O'Donnell LJ, Wells WM 3rd, Golby AJ, Westin CF. Unbiased groupwise registration of white matter tractography. *Med Image Comput Assist Interv.* 2012;15:123–130.
45. Zhang F, Wu Y, Norton I, *et al*. An anatomically curated fiber clustering white matter atlas for consistent white matter tract parcellation across the lifespan. *Neuroimage.* 2018;179:429–447.
46. Zhang F, Wu Y, Norton I, Rathi Y, Golby AJ, O'Donnell LJ. Test-retest reproducibility of white matter parcellation using diffusion MRI tractography fiber clustering. *Hum Brain Mapp.* 2019;40:3041–3057.
47. Wheeler AL, Voineskos AN. A review of structural neuroimaging in schizophrenia: from connectivity to connectomics. *Front Hum Neurosci.* 2014;8:653.
48. Kelly S, Jahanshad N, Zalesky A, *et al*. Widespread white matter microstructural differences in schizophrenia across 4322 individuals: results from the ENIGMA Schizophrenia DTI Working Group. *Mol Psychiatry.* 2018;23:1261–1269.
49. Hubert M, Vandervieren E. An adjusted boxplot for skewed distributions. *Comput Stat Data Anal.* 2008;52:5186–5201.
50. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. *J Stat Softw.* 2010;45:1–68.
51. Yeo I, Johnson RA. A new family of power transformations to improve normality or symmetry. *Biometrika.* 2000;87:954–959.
52. Elliott M. Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. In: Vittinghoff E, Glidden Dv, Shiboski Sc, Mcculloch Ce, eds. *Biometrics*. Vol. 62. Malden, USA: Blackwell Publishing; 2006:1271–1272.
53. Di Biase MA, Cropley VL, Cocchi L, *et al*. Linking cortical and connectional pathology in schizophrenia. *Schizophr Bull.* 2019;45:911–923.
54. Kanaan RA, Chaddock C, Allin M, *et al*. Gender influence on white matter microstructure: a tract-based spatial statistics analysis. *PLoS One.* 2014;9:e91109.
55. Monji A, Kato T, Kanba S. Cytokines and schizophrenia: microglia hypothesis of schizophrenia. *Psychiatry Clin Neurosci.* 2009;63:257–265.
56. Tabachnick BG, Fidell LS, Ullman JB. *Using Multivariate Statistics*. Vol. 5. Boston, MA: Pearson; 2007.
57. Zhuang X, Yang Z, Cordes D. A technical review of canonical correlation analysis for neuroscience applications. *Hum Brain Mapp.* 2020;41:3807–3833.
58. Thorndike RM, Weiss DJ. A study of the stability of canonical correlations and canonical components. *Educ Psychol Meas.* 1973;33:123–134.
59. Barcikowski RS, Stevens JP. A Monte Carlo study of the stability of canonical correlations, canonical weights and canonical variate-variable correlations. *Multivar Behav Res.* 1975;10:353–364.
60. Strand KH, Kossman S. Further inquiry into the stabilities of standardized and structure coefficients in canonical and discriminant analyses; 2000. *Online Submission*. Retrieved October 4, 2021, from http://files.eric.ed.gov/fulltext/ED572339.pdf.
61. Sherry A, Henson RK. Conducting and interpreting canonical correlation analysis in personality research: a user-friendly primer. *J Pers Assess.* 2005;84:37–48.
62. Dinga R, Schmaal L, Penninx BWJH, *et al*. Evaluating the evidence for biotypes of depression: Methodological replication and extension of Drysdale *et al*. (2017). *NeuroImage Clin.* 2019;22:101796.
63. Liu Y, Hayes DN, Nobel A, Marron JS. Statistical significance of clustering for high-dimension, low–sample size data. *J Am Stat Assoc.* 2008;103:1281–1293.
64. Adolphs R. The social brain: neural basis of social knowledge. *Annu Rev Psychol.* 2009;60:693–716.
65. Brunet-Gouet E, Decety J. Social brain dysfunctions in schizophrenia: a review of neuroimaging studies. *Psychiatry Res Neuroimaging*. 2006;148:75–92.
66. Pinkham AE, Penn DL, Perkins DO, Lieberman J. Implications for the neural basis of social cognition for the study of schizophrenia. *Am J Psychiatry.* 2003;160:815–824.
67. Van Overwalle F. Social cognition and the brain: a meta-analysis. *Hum Brain Mapp.* 2009;30:829–858.
68. Rocca P, Galderisi S, Rossi A, *et al*. Social cognition in people with schizophrenia: a cluster-analytic approach. *Psychol Med.* 2016;46:2717–2729.
69. Pinkham AE, Penn DL, Green MF, Buck B, Healey K, Harvey PD. The social cognition psychometric evaluation study: results of the expert survey and RAND panel. *Schizophr Bull.* 2014;40:813–823.
70. Nasrallah HA, Andreasen NC, Coffman JA, *et al*. A controlled magnetic resonance imaging study of corpus callosum thickness in schizophrenia. *Biol Psychiatry.* 1986;21:274–282.
71. Kochunov P, Coyle TR, Rowland LM, *et al*. Association of white matter with core cognitive deficits in patients with schizophrenia. *JAMA Psychiatry.* 2017;74:958–966.
72. Koshiyama D, Fukunaga M, Okada N, *et al*. Role of frontal white matter and corpus callosum on social function in schizophrenia. *Schizophr Res.* 2018;202:180–187.
73. Kubicki M, Westin C-F, Maier SE, *et al*. Uncinate fasciculus findings in schizophrenia: a magnetic resonance diffusion tensor imaging study. *Am J Psychiatry.* 2002;159:813–820.
74. Singh S, Singh K, Trivedi R, *et al*. Microstructural abnormalities of uncinate fasciculus as a function of impaired cognition in schizophrenia: a DTI study. *J Biosci.* 2016;41:419–426.
75. Jung S, Kim J-H, Sung G, *et al*. Uncinate fasciculus white matter connectivity related to impaired social perception and cross-sectional and longitudinal symptoms in patients with schizophrenia spectrum psychosis. *Neurosci Lett.* 2020;737:135144.
76. Kierońska S, Sokal P, Dura M, Jabłońska M, Rudaś M, Jabłońska R. Tractography-based analysis of morphological and anatomical characteristics of the uncinate fasciculus in human brains. *Brain Sci.* 2020:10. doi:10.3390/brainsci10100709
77. Hau J, Sarubbo S, Houde JC, *et al*. Revisiting the human uncinate fasciculus, its subcomponents and asymmetries with stem-based tractography and microdissection validation. *Brain Struct Funct.* 2017;222:1645–1662.
78. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med.* 2013;11:126.
79. Mikolas P, Hlinka J, Skoch A, *et al*. Machine learning classification of first-episode schizophrenia spectrum disorders and controls using whole brain white matter fractional anisotropy. *BMC Psychiatry.* 2018;18:97.
80. Pettersson-Yeo W, Benetti S, Marquand AF, *et al*. Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level. *Psychol Med.* 2013;43:2547–2562.

81. Liang S, Li Y, Zhang Z, *et al*. Classification of first-episode schizophrenia using multimodal brain features: a combined structural and diffusion imaging study. *Schizophr Bull.* 2019;45:591–599.
82. Deng Y, Hung KSY, Lui SSY, *et al*. Tractography-based classification in distinguishing patients with first-episode schizophrenia from healthy individuals. *Prog Neuropsychopharmacol Biol Psychiatry.* 2019;88:66–73.
83. Caan MWA, Vermeer KA, van Vliet LJ, *et al*. Shaving diffusion tensor images in discriminant analysis: a study into schizophrenia. *Med Image Anal.* 2006;10:841–849.
84. Chen Y-J, Liu C-M, Hsu Y-C, *et al*. Individualized prediction of schizophrenia based on the whole-brain pattern of altered white matter tract integrity. *Hum Brain Mapp.* 2018;39:575–587.
85. Caprihan A, Pearlson GD, Calhoun VD. Application of principal component analysis to distinguish patients with schizophrenia from healthy controls based on fractional anisotropy measurements. *Neuroimage.* 2008;42:675–682.
86. Rathi Y, Malcolm J, Michailovich O, *et al*. Biomarkers for identifying first-episode schizophrenia patients using diffusion weighted imaging. *Med Image Comput Comput Assist Interv.* 2010:657–665.
87. Ingalhalikar M, Kanterakis S, Gur R, Roberts TPL, Verma R. DTI based diagnostic prediction of a disease via pattern classification. *Med Image Comput Comput Assist Interv.* 2010;13:558–565.
88. Ardekani BA, Tabesh A, Sevy S, Robinson DG, Bilder RM, Szeszko PR. Diffusion tensor imaging reliably differentiates patients with schizophrenia from healthy volunteers. *Hum Brain Mapp.* 2011;32:1–9.
89. Clemm von Hohenberg C, Pasternak O, Kubicki M, *et al*. White matter microstructure in individuals at clinical high risk of psychosis: a whole-brain diffusion tensor imaging study. *Schizophr Bull.* 2013;40:895–903.
90. De Herdt A, Wampers M, Vancampfort D, *et al*. Neurocognition in clinical high risk young adults who did or did not convert to a first schizophrenic psychosis: a meta-analysis. *Schizophr Res.* 2013;149:48–55.
91. Piskulic D, Liu L, Cadenhead KS, *et al*. Social cognition over time in individuals at clinical high risk for psychosis: findings from the NAPLS-2 cohort. *Schizophr Res.* 2016;171:176–181.
92. Cropley VL, Klauser P, Lenroot RK, *et al*. Accelerated gray and white matter deterioration with age in schizophrenia. *Am J Psychiatry.* 2017;174:286–295.
93. Kurtz MM. Neurocognitive impairment across the lifespan in schizophrenia: an update. *Schizophr Res.* 2005;74:15–26.
94. Samartzis L, Dima D, Fusar-Poli P, Kyriakopoulos M. White matter alterations in early stages of schizophrenia: a systematic review of diffusion tensor imaging studies. *J Neuroimaging.* 2014;24:101–110.
95. Kanaan R, Barker G, Brammer M, *et al*. White matter microstructure in schizophrenia: effects of disorder, duration and medication. *Br J Psychiatry.* 2009;194:236–242.
96. Altamura AC, Serati M, Buoli M. Is duration of illness really influencing outcome in major psychoses? *Nord J Psychiatry.* 2015;69:1685403–1681699.
97. Szeszko PR, Robinson DG, Ikuta T, *et al*. White matter changes associated with antipsychotic treatment in first-episode psychosis. *Neuropsychopharmacology.* 2014;39:1324–1331.
98. Nielsen RE, Levander S, Kjaersdam Telléus G, Jensen SOW, Østergaard Christensen T, Leucht S. Second-generation antipsychotic effect on cognition in patients with schizophrenia--a meta-analysis of randomized clinical trials. *Acta Psychiatr Scand.* 2015;131:185–196.
99. Mendrek A, Mancini-Marïe A. Sex/gender differences in the brain and cognition in schizophrenia. *Neurosci Biobehav Rev.* 2016;67:57–78.
100. Shahab S, Stefanik L, Foussias G, Lai M-C, Anderson KK, Voineskos AN. Sex and diffusion tensor imaging of white matter in schizophrenia: a systematic review plus meta-analysis of the corpus callosum. *Schizophr Bull.* 2018;44:203–221.
101. Takahashi M, Iwamoto K, Fukatsu H, Naganawa S, Iidaka T, Ozaki N. White matter microstructure of the cingulum and cerebellar peduncle is related to sustained attention and working memory: a diffusion tensor imaging study. *Neurosci Lett.* 2010;477:72–76.
102. Cremers LGM, de Groot M, Hofman A, *et al*. Altered tract-specific white matter microstructure is related to poorer cognitive performance: the Rotterdam Study. *Neurobiol Aging.* 2016;39:108–117.
103. Helmer M, Warrington S, Mohammadi-Nejad AR, *et al*. On stability of Canonical Correlation Analysis and Partial Least Squares with application to brain-behavior associations. *BioRxiv.* 2020; doi:10.1101/2020.08.25.265546.

SUPPLEMENT S1: Visualization of selected white matter tracts



Note: AF, arcuate fasciculus; UF, uncinate fasciculus; ILF, inferior longitudinal fasciculus; IOFF, inferior occipital frontal fasciculus; TF, thalamo-frontal tract; CB, cingulum bundle; CC, corpus callosum; CC1, rostrum; CC2, genu; CC3, rostral body; CC4, anterior midbody; CC5, posterior midbody; CC6, isthmus; CC7, splenium.

SUPPLEMENT S2: Imaging quality control procedures

The SPINS study implemented several quality control procedures. To monitor and mitigate unwanted heterogeneity over five years of data collection, we used an agar phantom that allowed tracking of scanner drift (115), and annually scanned 'travelling human phantoms', allowing for cross-centre and longitudinal comparison of the same individuals as a proxy for inter- and intra-site scanner related impacts (116). The SPINS study also conducted centralized training for research assistants and implemented standard operating procedures across its three sites, that allowed for high quality diffusion images to be acquired from most study participants. In particular, participants were trained on how to minimise head motion, especially important given that the diffusion images were the last to be acquired. An in-scanner camera was employed to monitor participant movement during scans, and corrupted scans were repeated as possible.

Prior to analysis, all scans were checked for sufficient quality by experienced research staff, making use of an in-house quality control dashboard (https://github.com/TIGRLab/dashboard) that reviews both quantitative (e.g., framewise displacement, signal-to-noise) and qualitative metrics (e.g., detecting ghosting or blurring by eye). As a rule, we attempted to manually correct issues prior to pre-processing, as exclusion of 'noisy' data has been demonstrated to spuriously inflate group FA differences between individuals with schizophrenia and healthy controls (117). Ultimately, we removed data from one participant (SSD) on the basis of their acquired diffusion scan.

We also performed quality control of outputs from the whitematteranalysis pipeline. Specifically, we performed qualitative quality control at three points: (i) initial tractography, (ii) registration to the ORG atlas, and (iii) creation of the k=41 final tracts, on the the basis of macroscopic features (e.g. trajectory shape and volume). We removed data from four participants (all HC) on this basis. Notably, we achieved perfect inter-rater agreement on the plausibility of the trajectories (pass/fail) of final tracts in a subset of twenty participants, including two participants who were excluded (19 tracts x 20 participants = 380 comparisons). Quantitatively, we reviewed scalar values derived from the tracts for evidence of abnormal microstructural properties (i.e., outlying FA/MD/AD/RD values) and tract-level characteristics (e.g. outlying streamline count or length). We removed data from no participants on this basis.

SUPPLEMENT S3: Exploratory methods

In three exploratory analyses, we probed the utility of providing X'_1 variate scores to a clustering, cutpoint, and prediction analysis, in the spirit of data-driven nosology and precision modelling gaining traction in the literature. Because these scores are standardized (Z-scores), we combined scores computed separately within the discovery and validation sets, affording us a larger sample (n=308). We then split the combined discovery and validation samples into training (n=200, 117 SSD) and testing (n=108, 63 SSD) sets. This split is roughly equal to a 2:1 training:testing ratio (64.94%), and the size of the training set surpassed a widely-employed stability benchmark established for models containing neuroimaging features in schizophrenia (i.e., predictive models using structural imaging features in schizophrenia patients are not stable in sample sizes under n=130) (118). To avoid bias, the entirety of the three respective procedures is embedded within the validation framework (119).

Clustering analysis

The clustering analysis sought to determine if X'_1 scores might reveal innate brain-behaviour biotypes. We chose to perform hierarchical clustering, in part because hierarchical clustering can accommodate one-dimensional (only X'_1 is significant), whereas other algorithms (e.g., *k*-NN) require multidimensional space. First, we hierarchically clustered X'_1 scores using Ward's complete-linkage method (120), which we established to provide superior clustering structure between comparable agglomerative methods via highest coefficient (>.999) (121). Ward's method establishes clusters that minimize dispersion in Euclidean space, and has been usefully applied in the SPINS dataset(58,85). We set seven clusters as an arbitrary upper-limit (consistent with one cluster for each of the five SCID-IV-TR schizophrenia subtypes in our sample, one cluster for healthy controls, and one additional degree of freedom) and then determined the optimal number of clusters in this range via the Calinski-Harabasz index (CH, also known as the variance ratio criterion), which represents the ratio of between-cluster dispersion and inter-cluster dispersion (thus, a higher CH value indicates higher performance) (122). We tested if the observed CH_i index value was different from that derived from a null distribution of two-dimensional Gaussians with similar characteristics to X_i (mean, covariance, number of observations), but that embeds no underlying clusters by definition. We reasoned that indistinguishable CH indices would indicate our data likewise did not embed any true clusters(57). On the other hand, if an observed CH_i index surpassed a significant proportion of the null distribution, we could be confident in our identification of biologically distinct subgroups, and proceed to interrogate their meaningfulness in relation to psychopathological characteristics. We used the `cluster` (123) and `NbClust` (124) R packages for these analyses.

Cutpoint analysis

The cutpoint analysis aimed to determine if the continuous distribution of X'_1 scores might accurately dichotomize participants in accordance with their "ground-truth" diagnostic label. For this task, we split the combined discovery and validation samples (n=308) into training (n=200, 117 SSD) and testing (n=108, 63 SSD) sets, as for the prior clustering analysis. We stratified both samples by diagnostic group, to ensure an approximately equivalent ratio of participants with and without an SSD in the training and testing sets (~58.5%). Note that though the counts of those with and without an SSD differ within each set (reflective of the SPINS recruitment strategy preferential to participants with an SSD), this differences falls far short of "class imbalance", which refers to the case when one class is substantially underrepresented in the dataset(125).

To explore the diagnostic accuracy of X'_1 scores, we employed ROC (receiver operating characteristic) curve analysis(126), with X'_1 scores as the independent variable, and binary diagnostic label (i.e., SSD, HC) as the dependent variable. We fit a ROC curve in the training set, as well as a permuted training set, obtained by randomly shuffling known diagnostic labels (perm=1000). For both models, we calculated the AUC (area under the ROC curve) c-statistic, which provides a synthetic goodness-of-fit measure varying from 0 to 1 (AUC: 0.9–1.0=excellent; 0.8–0.9=good; 0.7–0.8=fair; 0.6–0.7=poor; 0.5–0.6=fail) (127), here indicating the extent to which X_i scores separate participants with and without SSD. We statistically

compared the AUC values from the training and permuted models via Chi-square, testing the null hypothesis that the vertical distance between their respective cumulative distribution functions did not differ from chance. This statistical test is essential before identifying a cutpoint: indeed, cutpoint analysis *will* identify a cutpoint in distributions with no diagnostic accuracy, just as clustering analyses will identify clusters in continuous data. The absence of this comparison has been cited as a potential cause of the poor performance and acceptance of purported biological cutpoints in clinical settings (128).

Contingent upon the diagnostic ability of X_i scores proving above chance, we then determined the 'optimal' location of a diagnostic cutpoint along the ROC curve. There exist several mathematical strategies to determine an 'optimal' cutpoint, that differ primarily in how they weigh the cost of misclassification(129). We opted to employ the Youden index (J) (130), which defines the optimal cutpoint as the maximum of the sum of sensitivity and specificity -1, with sensitivity and specificity afforded equal weight, and has been shown to derive the highest sensitivity estimates in cases when the 'non-diseased' population (here, HC) demonstrates high variability (131). Youden index values range from 0 to 1 (0.9-1.0=exceptional; 0.8–0.9=excellent; 0.7–0.8=acceptable; <0.7=no discrimination), and can be graphically represented as the longest vertical distance between the ROC curve and its 45 degree line of chance(132). We evaluated the classification performance of the derived Youden index, in both the training and testing sets, via several commonly-employed performance metrics (AUC, accuracy, sensitivity, specificity). We ensured that misclassifications of participants with SSD and HCs were equally likely, as determined by McNemar's test for marginal homogeneity. We analysed if the X'_1 scores exhibited qualitatively better classification performance (higher AUC values) than any of the 19 white matter variables informing the original CCA, as well as their one-dimensional representation achieved via PCA (95), by re-running the ROC analysis in the training set. We use the `rPROC` (133), `cutpointr` (134), and `caret` (135) R packages for these analyses.

Prediction analysis

Lastly, we sought to determine if X'_1 scores could predict social functioning, as measured by the Birchwood Social Functioning Scale (BSFS) (35). We chose to employ simple linear regression analysis rather than a more complex machine learning method, as the former retain interpretability and thus allow explanatory insights at the cost of minimally diminished performance (136). Clinically relevant prognosis and therapeutic discovery likely requires an adequate explanation of cause (137). BSFS total scores showed sufficient representation in the tails of the distribution, so we were able to treat functional outcome as a continuous index, which remains uncommon in the prognostic literature and is a consequence of assessing outcome across the healthy-to-schizophrenia spectrum (should this criterion have not been met, a classification model would be preferred to regression (138)). We employed 5-fold 5-repeat (internal) cross validation to determine adjusted R^2 goodness-of-fit (representing the proportion of variation in social functioning that is predicted by the model; higher is better) as well as RMSE prediction error (root mean squared error, representing the average prediction error made in predicting social functioning; lower is better). We opted to compare the goodness-of-fit of our model (X'_1 scores as the predictor variable) to an alternative model, using all 19 of the white matter FA features as predictor variables, to determine any 'value-added' by our cognition-constrained brain features. Though models with few predictors typically underfit the data and thus demonstrate lower variance explained than models with more predictors, additional predictors improve performance only if they provide meaningful information as opposed to noise (139). For this task, we implemented a likelihood ratio test modified for non-nested data (140). We used the `caret` (135) and `poweRlaw` (141) packages for these analyses.

SUPPLEMENT S4: Participant white matter fractional anisotropy estimates

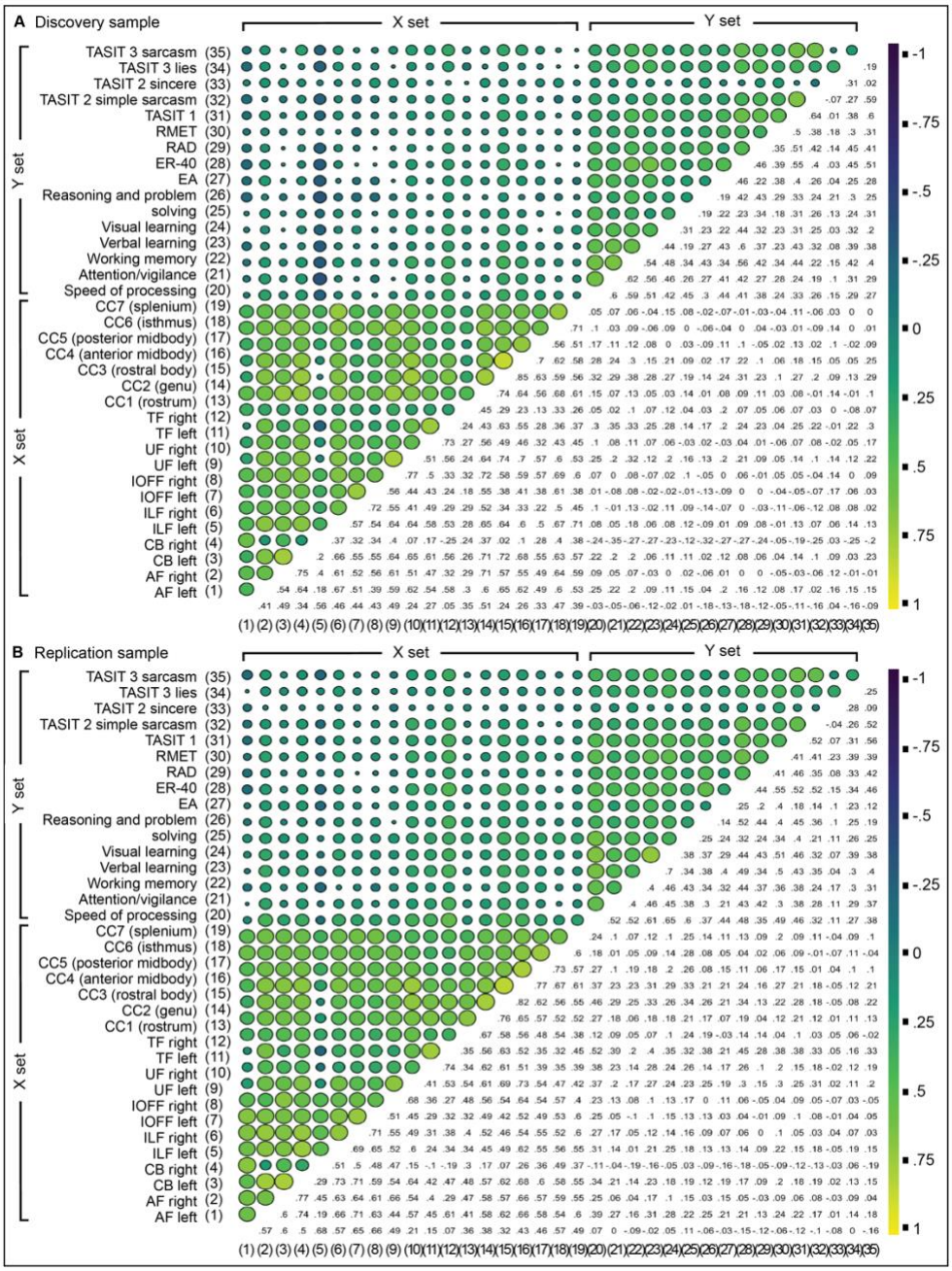
	Discovery sample (GE Discovery)				Validation sample (Siemens Prismas)				Sample comparison
	SSD-HC	SSD	HC		SSD-HC	SSD	HC		
	n=135	n=89	n=46		n=173	n=91	n=82		
	mean (SD)	mean (SD)	mean (SD)	p-value adj.†	mean (SD)	mean (SD)	mean (SD)	p-value adj.†	p-value adj. ‡
Fractional anisotropy									
AF left	0.60 (0.02)	0.60 (0.03)	0.61 (0.02)	1.000	0.53 (0.04)	0.52 (0.04)	0.53 (0.04)	0.274	0.000
AF right	0.57 (0.02)	0.57 (0.02)	0.57 (0.03)	1.000	0.52 (0.04)	0.51 (0.04)	0.53 (0.03)	0.041	0.000
CB left	0.50 (0.03)	0.50 (0.03)	0.50 (0.02)	1.000	0.44 (0.03)	0.43 (0.03)	0.44 (0.03)	0.102	0.000
CB right	0.49 (0.03)	0.48 (0.03)	0.49 (0.03)	1.000	0.43 (0.03)	0.42 (0.03)	0.43 (0.03)	0.044	0.000
ILF left	0.51 (0.02)	0.51 (0.02)	0.52 (0.02)	1.000	0.47 (0.03)	0.46 (0.03)	0.48 (0.03)	0.008	0.000
ILF right	0.51 (0.02)	0.51 (0.02)	0.52 (0.02)	1.000	0.47 (0.03)	0.46 (0.03)	0.48 (0.03)	0.004	0.000
IOFF left	0.63 (0.03)	0.63 (0.02)	0.63 (0.03)	1.000	0.54 (0.04)	0.53 (0.04)	0.55 (0.04)	0.015	0.000
IOFF right	0.63 (0.02)	0.62 (0.02)	0.63 (0.02)	1.000	0.55 (0.04)	0.54 (0.04)	0.55 (0.03)	0.021	0.000
TF left	0.49 (0.01)	0.49 (0.01)	0.50 (0.01)	1.000	0.46 (0.02)	0.45 (0.02)	0.46 (0.02)	0.158	0.000
TF right	0.48 (0.01)	0.48 (0.01)	0.49 (0.01)	0.720	0.46 (0.02)	0.46 (0.02)	0.46 (0.02)	0.111	0.000
UF left	0.48 (0.03)	0.49 (0.03)	0.48 (0.03)	1.000	0.43 (0.03)	0.43 (0.03)	0.44 (0.03)	1.000	0.000
UF right	0.46 (0.03)	0.46 (0.03)	0.46 (0.03)	1.000	0.44 (0.03)	0.44 (0.03)	0.44 (0.03)	1.000	0.000
CC1 (rostrum)	0.51 (0.03)	0.51 (0.03)	0.51 (0.02)	1.000	0.47 (0.03)	0.46 (0.03)	0.48 (0.02)	0.075	0.000
CC2 (genu)	0.57 (0.02)	0.57 (0.02)	0.57 (0.02)	1.000	0.53 (0.03)	0.53 (0.03)	0.54 (0.02)	0.007	0.000
CC3 (rostral body)	0.59 (0.02)	0.59 (0.02)	0.60 (0.02)	0.534	0.55 (0.03)	0.55 (0.03)	0.56 (0.02)	0.025	0.000
CC4 (anterior midbody)	0.60 (0.02)	0.60 (0.02)	0.60 (0.01)	0.593	0.57 (0.03)	0.56 (0.03)	0.57 (0.02)	0.063	0.000

CC5 (posterior midbody)	0.60 (0.02)	0.60 (0.02)	0.60 (0.02)	1.000	0.57 (0.03)	0.57 (0.03)	0.58 (0.02)	0.393	0.000
CC6 (isthmus)	0.61 (0.01)	0.61 (0.01)	0.62 (0.01)	0.408	0.57 (0.02)	0.57 (0.02)	0.58 (0.02)	0.037	0.000
CC7 (splenium)	0.61 (0.02)	0.61 (0.02)	0.62 (0.02)	1.000	0.58 (0.02)	0.58 (0.02)	0.59 (0.02)	0.047	0.000

All p-values are derived from a t-test. The three `p-value adj` columns have been Bonferroni corrected for multiple comparisons. † The p-values within the Discovery sample and Validation sample "blocks" respectively compare the SSD and HC scores (providing a comparison of participant groups within sample), and the respective SSD-HC columns describe the sample's combined participants' mean and standard deviation. ‡ The p-values in the Sample comparison block compare the combined SSD-HC means and standard deviations between the Discovery and Validation samples (providing a comparison across samples).

***Note*:** AF, arcuate fasciculus; CB, cingulum bundle; CC, corpus callosum; HC, healthy control; ILF, inferior longitudinal fasciculus; IOFF, inferior occipital-frontal fasciculus; SSD, schizophrenia spectrum disorder; SSD-HC, schizophrenia-to-healthy control spectrum; TF, ; UF, uncinate fasciculus.

SUPPLEMENT S5: Bivariate correlations of all CCA features



Bivariate correlations between all variables are shown in the (A) discovery sample and (B) replication sample. Correlation strength is denoted by colour and size in the upper triangle, and numerically in the lower triangle. Correlations were highest within the X set (mean=.52, min=-.26, max=.85), followed my within the Y set (mean=.35, min=-.12, max=.72), and lastly between sets (mean=.06, min=-.35, max=.37).