**Supplementary material**

[Supplement 1: Patients’ medication and comorbidities 2](#_Toc120633831)

[Supplement 2: Details on cognitive tests 3](#_Toc120633832)

[Verbal learning and memory test (VLMT) 3](#_Toc120633833)

[Digit symbol substitution test (DSST) 3](#_Toc120633834)

[Trail-Making-Test A and B (TMT-A & TMT-B) 3](#_Toc120633835)

[d2 Test of Attention (d2) 4](#_Toc120633836)

[Letter number sequencing test (LNST) 4](#_Toc120633837)

[Corsi Block-Tapping Task 4](#_Toc120633838)

[Regensburger Wortflüssigkeitstest (RWT) for verbal fluency 4](#_Toc120633839)

[Supplement 3: MRI data acquisition 5](#_Toc120633840)

[Supplement 4: Preprocessing of diffusion-weighted images 5](#_Toc120633841)

[Supplement 5: Anatomical connectome reconstruction 5](#_Toc120633842)

[Supplement 6: Quality control procedure for connectivity matrices 6](#_Toc120633843)

[Supplement 7: Calculation of standard graph metrics 7](#_Toc120633844)

[Supplement 8: Python packages employed within the analyses 7](#_Toc120633845)

[Supplement 9: Details on exploratory factor analysis (EFA) 8](#_Toc120633846)

[Data preparation 8](#_Toc120633847)

[Assessment of adequacy 8](#_Toc120633848)

[Determination of the number of factors to be extracted 8](#_Toc120633849)

[Extraction of factor scores 11](#_Toc120633850)

[Supplement 10: Details on the NBS method 13](#_Toc120633851)

[Supplement 11: Subnetwork-specific analyses based on an NBS threshold of *F =* 4.0 14](#_Toc120633852)

[Supplement 12: Relationship between NBS F-threshold and network size 15](#_Toc120633853)

[Supplement 13: Results from exploratory mediation analysis 16](#_Toc120633854)

[Supplement 14: Tests statistics representing between-group differences in cognitive performance 17](#_Toc120633855)

[Supplement 15: Differences in CF-VF when controlling for clinical covariates 18](#_Toc120633856)

[Supplement 16: Between-group differences in cognitive performance due to comorbidity 18](#_Toc120633857)

[Supplement 17: Between-group differences in cognitive performance when excluding outliers 18](#_Toc120633858)

[Supplement 18: Remission status-related interaction effects 20](#_Toc120633859)

[Interaction effects on global-level cognition-connectome associations 20](#_Toc120633860)

[Interaction effects on local-level cognition-connectome associations 20](#_Toc120633861)

[Supplement 19: Description of identified networks 21](#_Toc120633862)

[Overlap of network edges 21](#_Toc120633863)

[Patterns of structural connectivity between brain regions 21](#_Toc120633864)

[Involvement of left and right hemispheres 23](#_Toc120633865)

[Supplement 20: Robustness checks of NBS analyses 25](#_Toc120633866)

[Supplement 21: Full results from correlational analyses 26](#_Toc120633867)

[Supplement 22: Symptom specific partial correlations 26](#_Toc120633868)

[Supplement 23: Correlational analyses correcting for clinical covariates 27](#_Toc120633869)

[Supplementary references 29](#_Toc120633870)

# Supplement 1: Patients’ medication and comorbidities

|  |
| --- |
| **Table S1.T1.***Details on patients’ medication.* |
| **Medication** | **MDDa (n = 307)** | **MDDr (n = 372)** | $$χ²$$ | **p** |
| Antidepressants: |  |  |  |  |
| SSRI | 68 | 54 | 4.107 | 0.043 |
| SNRI | 70 | 49 | 7.713 | 0.006 |
| NaSSA | 30 | 18 | 4.584 | 0.032 |
| NDRI | 6 | 4 | 0.600 | 0.439 |
| NaRI | 2 | 0 | 2.185 | 0.139 |
| TCA | 21 | 8 | 7.459 | 0.006 |
| MAOI | 0 | 1 | 0.922 | 0.337 |
| Agomelatine | 8 | 11 | 0.267 | 0.605 |
| Antipsychotics | 54 | 26 | 14.952 | 0.000 |
| Lithium | 5 | 2 | 1.577 | 0.209 |
| **Note.** The table compares the frequency of medications taken from different drug classes between acutely depressed (MDDa) and remitted (MDDr) patients with major depressive disorder in our sample. SSRI = selective serotonin reuptake inhibitor; SNRI = selective serotine-norepinephrine reuptake inhibitor, NaSSA = noradrenergic and specific serotonergic antidepressant; NDRI = norepinephrine-dopamine reuptake inhibitor, NaRI = noradrenaline reuptake inhibitor, TCA = tricyclic antidepressants, MAOI = monoamine oxidase inhibitors;  |

|  |
| --- |
| **Table S1.T2.***Details on patients’ comorbidities.* |
| **Diagnosis** | **MDDa (n = 307)** | **MDDr (n = 372)** |
| Adjustment disorder | 0 | 1 |
| Anxiety disorder | 149 | 145 |
| Dysthymia | 25 | 9 |
| Eating disorder | 31 | 30 |
| Impulse-Control Disorders Not Elsewhere Classified | 1 | 0 |
| Brief psychotic disorder | 2 | 2 |
| Psychotic Disorder NOS | 0 | 1 |
| Delusional disorder | 0 | 2 |
| Somatoform disorders | 12 | 14 |
| Substance-related disorder | 18 | 16 |
| **Note.** The table compares the frequencies of (lifetime) comorbid psychiatric disorders between patients diagnosed with acute (MDDa) or remitted (MDDr) major depressive disorder. |

# Supplement 2: Details on cognitive tests

We included scores from the following cognitive tests from the neuropsychological test battery of the Marburg-Münster Affective Disorders Cohort Study in our analysis to assess the participants’ cognitive performance:

## Verbal learning and memory test (VLMT)

To assess the participants’ ability to learn, recall and recognize verbal information, we included scores from the Verbal Learning and Memory Test (“Verbaler Lern- und Merkfähigkeitstest”, VLMT) (1) in our analysis, which is a German version of the Rey Auditory Verbal Learning Test (RAVLT). The test includes 1) serial learning of word lists (test score VLMT 1-5), defined as the total number of correctly remembered words during the five list learning iterations; 2) retrieval of the learned words after learning an interference list (VLMT 6), defined as the number of correctly remembered words from the first list after learning the interference list; 3) retrieval of words after a delay of 20-25 minutes (VLMT 7), defined as the number of correctly remembered words after the delay; and 3) recognition of words from a series of distractor words (VLMT 8), defined as the number of correctly recognized words from the first list.

## Digit symbol substitution test (DSST)

The Digit symbol substitution test (2) is a subtest of the Wechsler Adult Intelligence Scale – Revised (3) and is considered a measure of processing speed. In this test, the participant has to assign the correct symbol to as many digits as possible within 90 seconds based on a predefined key. The score is the total number of correct assigned symbols.

## Trail-Making-Test A and B (TMT-A & TMT-B)

We included both subtests from the Trail-Making Test (4) in our analyses. The first subtest (TMT-A) consists of 25 circles containing a number between 1 and 25. The participant must draw a line as quickly as possible that connects all the circles in ascending order of their numbers. The time the participant needs to complete the task is the score in this test. This score is considered a measure of processing speed and visual scanning. In the second subtest (TMT-B), the participant must draw a line connecting circles as quickly as possible. However, besides circles containing numbers, the TMT-B also includes circles containing letters from A to L. As with TMT-A, in TMT-B, the participant must connect the circles in ascending order. To this end, the participant must alternate between circles with numbers and circles with letters, i.e., the participant begins with 1, then draws a line to A, and then a line to 2, and so on. We used the score of this subtest, which is the time needed to complete the task, as a measure of cognitive flexibility and set-shifting.

## d2 Test of Attention (d2)

The d2 Test of Attention (5) is considered a measure of selective and sustained attention. In this task, the participant must scan 14 lines of 47 letters ("d" or "p") each. The letters are each marked with one, two, three, or four dashes. The participant has to cross out only those d's that have precisely two dashes. For each line, the participant is given 20 seconds. Concentration performance is determined as the number of letters crossed out minus letters missed or incorrect.

## Letter number sequencing test (LNST)

The Letter number sequencing test (LNST) is a subtest of the Wechsler Adult Intelligence Scale – Revised (3) is considered a measure of verbal working memory capacity. In each trial of the test, a sequence of numbers and letters is read to the participant. The participant must then mentally sort this sequence by naming all the numbers in ascending order and then all the letters in alphabetical order. Throughout the task, the sequences become increasingly longer. The test result is defined as the number of sequences correctly sorted by the participant.

## Corsi Block-Tapping Task

The Corsi Block-Tapping Task (6) is considered a measure of visuospatial short-term and working memory. In this test, the participant is presented with a board on which nine blocks are placed in an imbalanced arrangement. In the first subtest (Corsi Block forward), the examiner taps on a sequence of these blocks. The participant must memorize this sequence and then tap the blocks in the same order. In the second subtest (Corsi block backward), the examiner again taps on a sequence of blocks, but this time the participant must repeat the sequence in reverse order, i.e., he or she begins with the block that the examiner last tapped. In both subtests, the sequences become longer throughout the task. In both subtests, the result is defined as the number of correctly (re-)produced sequences.

## Regensburger Wortflüssigkeitstest (RWT) for verbal fluency

To assess the participants’ verbal fluency, we included three scores from the Regensburger Wortflüssigkeitstest (RWT) (7). The first subtest (VF Letter) measures phonematic word fluency by asking the participant to name as many words as possible, beginning with the letter p within one minute. The second subtest (VF category) assesses semantic word fluency by asking the participant to name as many words as possible that belong to the category animals. The third subtest (VF category alternating) assesses cognitive flexibility and semantic word fluency by asking the participant to name words that alternately belong to one of the two categories *sports* and *fruits*. The number of correctly named words within one minute is used as the final score for all tests.

# Supplement 3: MRI data acquisition

T1 and diffusion-weighted images (DWI) were acquired using 3T whole-body MRI scanners (Marburg: Tim Trio, 12-channel head matrix Rx-coil, Siemens, Erlangen, Germany; Münster: Prisma, 20-channel head matrix Rx-coil, Siemens, Erlangen, Germany). At both scanners, a GRAPPA acceleration factor of two was applied for both sequences. A high resolution T1-weighted dataset were acquired using a 3D-MPRAGE-sequence (TE = 2.28 ms, TR = 2130 ms, TI = 900 ms) with an isotropic voxel size of 1 x 1 x 1 mm³. For DTI imaging, fifty-six axial slices with no gap, were measured with an isotropic voxel size of 2.5 x 2.5 x 2.5 mm³ (TE=90 ms, TR=7300 ms). Five non-DW images (b=0 s/mm²) and 2 x 30 DW images with a b-value of 1000 s/mm² were acquired. Imaging pulse sequence parameters were standardized across both sites to the extent permitted by each platform.

# Supplement 4: Preprocessing of diffusion-weighted images

Diffusion-weighted images (DWI) were realigned and corrected for eddy currents and susceptibility distortions (8) using the eddy function from FSL 6.0.1 (9,10). Diffusion tensor imaging models the measured signal of a voxel by a single tensor describing the diffusion signal as one preferred diffusion direction per voxel. The CATO toolbox (11) employed for reconstruction of the anatomical connectome (see Supplement 6) uses the informed RESTORE algorithm (12,13) that estimates the tensor while identifying and removing outliers during the fitting, thereby reducing the impact of physiological noise artifacts on the DTI modeling. Based on the diffusion profiles, white matter pathways were reconstructed using deterministic tractography. To this end, eight seeds were started per voxel, and for each seed, a tractography streamline was constructed by following the main diffusion direction from voxel to voxel. Stop criteria included reaching a voxel with a fractional anisotropy < 0.1, making a sharp turn of >45°, reaching a gray matter voxel, or exiting the brain mask (14).

# Supplement 5: Anatomical connectome reconstruction

We employed the publicly available [CATO toolbox](http://www.dutchconnectomelab.nl/CATO/) (11) for reconstructing the anatomical connectome. The procedure included the following steps:

We obtained a network of 114 brain regions along with the reconstructed white matter streamlines between these brain areas for each participant. To identify the brain areas, we relied on FreeSurfer's Desikan-Killiany Atlas (15–17). However, given the poorer DWI signal-to-noise ratio in subcortical regions and the dominant effect of subcortical regions on network properties, we decided to use a subdivision of this atlas containing only cortical regions, as we have done in previous work (18,19).

To reconstruct the streamlines, we applied a deterministic streamline tractography based on the Fiber Assignment by Continuous Tracking (FACT) algorithm (20). We chose this deterministic algorithm instead of more advanced diffusion direction reconstruction methods because it provides a reasonable balance between false-negative and false-positive fiber reconstructions (21). Connections between two nodes, i.e., brain areas, were included if at least three reconstructed streamlines connected them. This type of thresholding was applied since we wanted to balance the sensitivity and specificity of the resulting connectivity matrices (22,23). In addition, given previous findings showing strong age-cognition (24–26) and age-structural connectivity associations (27–29), age is likely to affect cognition-structural connectivity associations. Since a recent analysis found thresholded connectivity matrices to provide more age-sensitive network measures (30), we expected that thresholding the matrices would improve accuracy when detecting age-related effects on structural connectivity, which in turn would improve our attempts to control for this confounding variable in our analyses.

Each participant's network was finally stored in a connectivity matrix with rows and columns representing nodes and matrix entries representing edges, i.e., connectivity strength measured as the number of reconstructed streamlines (NOS) between two nodes.

# Supplement 6: Quality control procedure for connectivity matrices

To ensure the quality of the connectivity matrices, we followed the approach from (31) and applied several criteria to identify outliers within the matrices. The criteria for outlier detection included (1) the average number of streamlines, (2) the average fractional anisotropy, (3) the average prevalence of each participant's connections (low value if the participant has "odd" connections), and (4) the average prevalence of each participant's connected brain regions (high value if the participant misses frequent connections). For each metric, quartiles (Q1, Q2, Q3) and interquartile range (IQR=Q3-Q1) were calculated. A data point was declared an outlier if its value was below Q1-1.5\*IQR or above Q3+1.5\*IQR for any of the four metrics. Application of this quality control procedure led to the exclusion of 15 MDD and 12 HC in the Marburg sample and 18 MDD and 15 HC in the Münster sample.

# Supplement 7: Calculation of standard graph metrics

Based on binarized connectivity matrices, we calculated the following graph metrics to assess the topological organization of the anatomical connectome (32):

|  |
| --- |
| **Table S7.T1***Definition and interpretation of graph metrics employed in our analyses.* |
| **Metric** | **Definition** | **Interpretation** |
| Shortest path length (L) | Average shortest path length between all node pairs | functional integration of the connectome |
| Global efficiency (GE) | Inverse of L | overall communication capacity of the connectome |
| Clustering coefficient (C) | Average probability that a node’s neighbors are also connected | functional segregation of the connectome |
| Normalized shortest path length (Lnormalized) | Normalized1 L | Like L, but corrected for differences in numbers of edges between participants |
| Normalized global efficiency (GEnormalized) | Normalized1 GE | Like GE, but corrected for differences in numbers of edges between participants |
| Normalized Clustering coefficient (Cnormalized) | Normalized1 C | Like C, but corrected for differences in numbers of edges between participants |
| Small-world index (SW) | GEnormalized/Lnormalized | Extend to which a network is more clustered than a random network while having a comparable shortest path length |
| **Note.** The table shows definitions and interpretations of all graph metrics employed in our analyses to assess the topological organization of anatomical brain networks.1To account for differences in the number of edges between participants, we calculated normalized equivalents of L. GE and C. Therefore, 1,000 random networks were generated based on each participant's connectivity matrix. Then, the normalized measures were calculated as the ratio between the measure and the average measure of the random networks.  |

# Supplement 8: Python packages employed within the analyses

|  |
| --- |
| **Table S8.T1***Python packages employed within the analyses.* |
| **Package** | **Version** | **Reference** |
| factor\_analyzer | 0.3.2.1 | (33) |
| matplotlib | 3.3.1 | (34) |
| numpy | 1.19.1 | (35) |
| pandas | 1.2.4 | (36) |
| pingouin | 0.3.12 | (37) |
| researchpy | 0.2.3 | (38) |
| scipy | 1.6.2 | (39) |
| seaborn | 0.10.1 | (40) |
| statsmodels | 0.12.2 | (41) |

# Supplement 9: Details on exploratory factor analysis (EFA)

The exploratory factor analysis (EFA) was carried out based on the recommendations in (42). Our procedure can be divided into four steps: 1) data preparation, 2) assessment of the adequacy of conducting exploratory factor analysis, 3) determining the number of factors to extract, and 4) extraction of the factor scores. If not stated otherwise, calculations carried out in steps 2 - 4 were conducted using Python’s factor\_analyzer package (33).

## Data preparation

A total of 14 cognitive test scores (see Supplement 10) were included in the EFA. Scores from TMT-A and TMT-B were inverted by multiplying them with -1 to ensure that higher values indicate higher cognitive performance in all tests. All variables were z-standardized before being entered into the analysis.

## Assessment of adequacy

To assess the adequacy of conducting exploratory factor analysis, we relied on Bartlett’s test and Kaiser-Meyer-Olkin (KMO) test results.

Bartlett’s test checks whether the observed variables are intercorrelated by testing the observed correlation matrix against an identity matrix. The test was significant (𝛘2 = 10567.969, *p* < 0.001), suggesting intercorrelation between the variables.

The KMO test establishes the suitability of data for factor analysis by estimating the proportion of shared variance among all observed variables. KMO values range between 0 and 1. Values of KMO less than 0.6 are considered inadequate. In our analysis, the value was KMO = .9, indicating that a sufficiently large proportion of the variance is shared among the variables.

Taken together, both tests confirm that EFA can be carried out based on our dataset.

## Determination of the number of factors to be extracted

The number of factors to be extracted was determined using the Kaiser criterion, the Scree plot, and by conducting a parallel analysis (43).

The Kaiser criterion is based on the eigenvalues of the factors. A factor’s eigenvalue is defined as the proportion of variance from all observed standardized variables explained by that factor. It is obtained by adding the squared factor loadings across all variables. Since standardized variables have a variance of 1, an eigenvalue > 1 indicates that this factor explains more variance than one variable alone. If so, this factor can be used to reduce the data. The Kaiser criterion relies on this assumption, i.e., according to this criterion, all factors with an eigenvalue > 1 are considered relevant. In our analysis, the eigenvalues of three factors were > 1 (see Table S9.T1 below), suggesting to extract three factors.

|  |
| --- |
| **Table S9.T1***Eigenvalues of all factors from exploratory factor analysis* |
| **Factor:** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| **Eigenvalue:** | 5.769 | 1.752 | 1.234 | 0.887 | 0.712 | 0.64 | 0.548 | 0.52 | 0.506 | 0.414 | 0.37 | 0.351 | 0.194 | 0.104 |
| **Note.** The table shows the eigenvalues of each factor that could be extracted from the data of cognitive tests. |

The scree plot is obtained by plotting each factor’s eigenvalue as a function of the number of that factor. All factors whose eigenvalues in the scree plot lie before the bend of the line are considered relevant. However, in our analysis, the interpretation of the scree plot (see Figure S9.F1) was not clear. Based on the plot, both the extraction of one factor and the extraction of 3 factors would have been adequate.

|  |
| --- |
| **Figure S9.F1***Scree plot for data from cognitive tests.* |
| C:\Users\m_grub03\owncloud\TraP\Projektordner\FactorAnalysis_Cognition_Connectome\3_Results\results\ScreePlot.png |

The parallel analysis (43) addresses the problem that eigenvalues > 1 can be based on random correlations given in the data but are not observable in the population that these data are supposed to represent. For this purpose, the observed eigenvalues are compared with eigenvalues resulting from a parallel analysis with variables that are uncorrelated (orthogonal) in the population but have nonzero random correlations in the sample. For our analyses, we translated the parallel analysis based on the fa.parallel function implemented in the R package psych (44,45) into Python. The algorithm performs a 1000-fold permutation of the data used for factor analysis and extracts the eigenvalues of all possible factors at each iteration. Then, the eigenvalues found based on the actual data set are compared to the eigenvalues found based on the permuted data sets. The significance of the factors is calculated as the number of iterations in which the eigenvalue of a factor calculated based on the actual data set was smaller than the eigenvalues of the identical factor calculated based on the permuted data sets, divided by the number of iterations. The result was tested against the results provided by the R function fa.parallel.

In our analysis, parallel analysis (see Table S9.T2 and Figure S9.F2) suggests extracting three factors. Since this result agrees with evaluating the eigenvalues using the Kaiser criterion and because the scree plot does not contradict this judgment, three factors were extracted as the result of the EFA.

|  |
| --- |
| **Table S9.T2***Results from the parallel analysis.* |
|  **Factor** | **Eigenvalue** | **Proportion of variance explained** | **LB-CI of Eigenvalues from parallel analysis** | **UB-CI of Eigenvalues from parallel analysis** | **p-value** |
| 1 | 5.773 | 0.412 | 1.135 | 1.199 | 0.0 |
| 2 | 1.753 | 0.125 | 1.104 | 1.153 | 0.0 |
| 3 | 1.235 | 0.088 | 1.079 | 1.120 | 0.0 |
| 4 | 0.887 | 0.063 | 1.056 | 1.093 | 1.0 |
| 5 | 0.712 | 0.051 | 1.034 | 1.069 | 1.0 |
| 6 | 0.640 | 0.046 | 1.012 | 1.047 | 1.0 |
| 7 | 0.548 | 0.039 | 0.993 | 1.024 | 1.0 |
| 8 | 0.520 | 0.037 | 0.970 | 1.003 | 1.0 |
| 9 | 0.506 | 0.036 | 0.951 | 0.984 | 1.0 |
| 10 | 0.415 | 0.030 | 0.929 | 0.965 | 1.0 |
| 11 | 0.371 | 0.026 | 0.907 | 0.943 | 1.0 |
| 12 | 0.351 | 0.025 | 0.883 | 0.921 | 1.0 |
| 13 | 0.194 | 0.014 | 0.857 | 0.899 | 1.0 |
| 14 | 0.104 | 0.007 | 0.818 | 0.873 | 1.0 |
| **Note.** The table shows the results from the parallel analyses conducted to determine the number of factors to extract from the dataset. LB-CI = Lower bound of the 95%-confidence interval around the eigenvalues established in the parallel analyses, UB-CI = Upper bound of the 95%-confidence interval around the eigenvalues established in the parallel analyses. |

|  |
| --- |
| **Figure S9.F2***Results from the parallel analysis.* |
| C:\Users\m_grub03\owncloud\TraP\Projektordner\FactorAnalysis_Cognition_Connectome\3_Results\results\ParallelAnalysis.png |
| **Note.** The figure shows the results from the parallel analyses conducted to determine the number of factors to extract from the dataset. The blue line represents the eigenvalues calcualted for each factor based on the actual dataset. The green and orange lines represent the upper and lower bound of the 95%-confindence interval of the eigenvalues calculated based on the permuted data sets. |

## Extraction of factor scores

We performed an exploratory factor analysis (EFA) to extract the factor scores. That EFA extracted three factors using the oblimin method as rotation method. We opted for this non-orthogonal instead of an orthogonal rotation method since the resulting factor scores, which all assess cognitive performance, likely would not be independent of each other, even though they capture different cognitive domains. The resulting loadings of the cognitive test scores on the three extracted factors are shown in Table S9.T3 below. Since especially measures of processing speed (DSST, TMT-A), attention (D2), and lower executive functions (set-shifting (TMT-B), visual working memory (block span forward and backward), verbal working memory (LNST)) loaded on the first factor, this factor was labeled as processing speed/lower executive functions factor (CF-PS). Because the second factor was almost exclusively characterized by the Verbal learning and memory test (VLMT), that factor was interpreted as the verbal learning and memory factor (CF-VLM). Since measures of phonematic and semantic verbal fluency loaded on the third factor, that factor was interpreted as the verbal fluency factor (CF-VF). Note, however, that naming the factors was done to facilitate referencing the factors. Thus, the pattern of factor loadings should always be taken into account when reading the results.

|  |
| --- |
| **Table S9.T3** *Factor loadings of the cognitive tests.* |
| **Test score** | **Cognitive domain** | **Factor loadings of the cognitive tests** |
| **CF-PS** | **CF-VLM** | **CF-VF** |
| VLMT 1-5 | Verbal learning | 0.077 | 0.839 | 0.078 |
| VLMT 6 | Verbal memory: Free recall after interference | -0.010 | 0.937 | 0.011 |
| VLMT 7 | Verbal memory: Delayed free recall | -0.013 | 0.953 | 0.002 |
| VLMT 8 | Verbal memory: Recognition | -0.027 | 0.858 | -0.075 |
| DSST | Processing speed | 0.702 | 0.012 | 0.174 |
| TMT-A | Processing speed, visual scanning | 0.757 | -0.080 | 0.043 |
| D2 concentration | Selective attention | 0.678 | 0.061 | 0.118 |
| LNST | Verbal working memory | 0.450 | 0.139 | 0.135 |
| TMT-B | Cognitive flexibility, set-shifting | 0.711 | 0.012 | 0.116 |
| Corsi block fwd | Visuospatial memory span | 0.739 | 0.002 | -0.196 |
| Corsi block bw. | Visuospatial working memory | 0.718 | 0.088 | -0.162 |
| VF letter | Phonemic verbal fluency | 0.024 | 0.024 | 0.764 |
| VF category | Semantic verbal fluency | -0.033 | 0.045 | 0.728 |
| VF category altern. | Cognitive flexibility, semantic verbal fluency | 0.070 | -0.005 | 0.766 |
| **Note.** CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance, VLMT = Verbal learning and memory test, DSST = Digit symbol substitution test, TMT = Trail making test, LNST = Letter number sequencing test, Corsi block fwd = Corsi block tapping test forward, Corsi block bw. = Corsi block tapping test backwards, VF = Verbal fluency, altern = alternating |

# Supplement 10: Details on the NBS method

We employed Network-based statistics (NBS, (46)) to identify networks of edges related to a given cognitive factor. NBS identifies a cluster-level effect by performing univariate mass tests at the edge level, controlling for family-wise errors (FWE). We used sex, study site, and age as covariates. The analysis proceeded as follows: Each edge was first assigned an *F*-value reflecting the association between the NOS of that edge and the respective cognitive factor. The *F*-values were thresholded at *F =* 5.8[[1]](#footnote-1) to select all suprathreshold edges. The largest network of suprathreshold edges was then selected to identify the most robust network of edges associated with the respective cognitive factor. This procedure was then repeated 5000 times in permutation tests (randomly assigning the cognitive factor values after regressing out the covariates) to identify an empirical null distribution describing the size of networks (47). The resulting p-value indicates the proportion of permutations in which the largest network in the null distribution was larger than the one initially identified. Note that we set the significance level to ɑ=.05/number of NBS analyses=.05/3 cognitive factors=.0167 in these analyses to correct for multiple testing.

In a second group of NBS analyses, we tested the remission status x cognitive performance interaction effects on the structural connectome while correcting for age, sex, and study-site. In these analyses, each edge was assigned an *F*-value reflecting its association with the remission status x cognitive factor interaction effects. The *F*-values were again thresholded at *F* = 5.8 in to obtain results at threshold levels in line with the above main effects analyses. The largest network of suprathreshold edges was then selected to identify the most robust network of edges associated with the interaction effect of the remission status and the respective cognitive factor. This procedure was repeated 5000 times in permutation tests (randomly assigning the remission status x cognitive factor interaction effect values after regressing out the covariates) to identify an empirical null distribution describing the size of networks. The resulting p-value indicates the proportion of permutations in which the largest network in the null distribution was larger than the one initially identified. In these analyses, we again set the significance level to ɑ=.05/number of NBS analyses=.05/3 cognitive factors=.0167 to correct for multiple testing.

# Supplement 11: Subnetwork-specific analyses based on an NBS threshold of *F =* 4.0

To evaluate the impact of the F-threshold chosen when employing the NBS toolbox, we repeated our analyses based on a network that was found when using an *F-*threshold of *F =* 4.0 instead of *F =* 5.8. This resulted in considerably less specific networks: Compared to the networks identified at a threshold of *F* = 5.8, the networks at a threshold of *F* = 4.0 contained up to three times as many edges (CF-PS network: +144%; CF-VLM network: +91%; CF-VF network: +226%). Considering these growth rates, the application of a low threshold has a relatively small impact on the overall pattern of results. We still find a reduced connectivity strength in MDDa compared to HC within the CF-PS related subnetwork (Table S11.T1) as well as correlations between connectivity strength within the CF-PS network, current depressive symptom severity, and performance on the CF-PS factor (Table S11.T2), although we acknowledge reduced effect sizes in both analyses.

|  |
| --- |
| **Table S11.T1***Group differences in sum of NOS within identified networks at NBS F-threshold of F = 5.8 and F = 4.0 - sample grouped by remission status* |
| **Network** | **Contrast** | **Mean & SD Group A** | **Mean & SD Group B** | **T** | **p** | **PFDR** | **Cohen's d** |
| *Group differences when applying an NBS t-threshold of F = 5.8* |
| **CF-PS network**(77 edges) | HC vs. MDDa | 38.6 ± 8.0 | 37.3 ± 7.6 | 3.032 | 0.003 | 0.008 | 0.200 |
| HC vs. MDDr | 38.6 ± 8.0 | 38.2 ± 8.4 | 1.290 | 0.198 | 0.198 | 0.082 |
| MDDa vs. MDDr | 37.3 ± 7.6 | 38.2 ± 8.4 | -1.497 | 0.135 | 0.198 | -0.115 |
| **CF-VLM network**(105 edges) | HC vs. MDDa | 12.9 ± 3.0 | 12.9 ± 3.1 | 0.479 | 0.632 | 0.956 | 0.032 |
| HC vs. MDDr | 12.9 ± 3.0 | 12.9 ± 3.0 | 0.460 | 0.645 | 0.956 | 0.029 |
| MDDa vs. MDDr | 12.9 ± 3.1 | 12.9 ± 3.0 | -0.056 | 0.956 | 0.956 | -0.004 |
| **CF-VF network**(65 edges) | HC vs. MDDa | 22.4 ± 5.3 | 21.5 ± 5.2 | 1.686 | 0.092 | 0.277 | 0.112 |
| HC vs. MDDr | 22.4 ± 5.3 | 21.6 ± 4.9 | 0.934 | 0.351 | 0.459 | 0.057 |
| MDDa vs. MDDr | 21.5 ± 5.2 | 21.6 ± 4.9 | -0.740 | 0.459 | 0.459 | -0.057 |
| *Group differences when applying an NBS t-threshold of F = 4.0* |
| **CF-PS network**(188 edges) | HC vs. MDDa | 24.2 ± 4.4 | 23.8 ± 4.2 | 2.101 | 0.036 | 0.108 | 0.140 |
| HC vs. MDDr | 24.2 ± 4.4 | 24.0 ± 4.5 | 1.149 | 0.251 | 0.376 | 0.073 |
| MDDa vs. MDDr | 23.8 ± 4.2 | 24.0 ± 4.5 | -0.848 | 0.397 | 0.397 | -0.065 |
| **CF-VLM network**(201 edges) | HC vs. MDDa | 12.0 ± 2.4 | 11.9 ± 2.5 | 0.835 | 0.404 | 0.738 | 0.057 |
| HC vs. MDDr | 12.0 ± 2.4 | 12.0 ± 2.5 | 0.335 | 0.738 | 0.738 | 0.021 |
| MDDa vs. MDDr | 11.9 ± 2.5 | 12.0 ± 2.5 | -0.469 | 0.639 | 0.738 | -0.036 |
| **CF-VF network**(212 edges) | HC vs. MDDa | 15.6 ± 3.0 | 15.2 ± 3.2 | 1.425 | 0.155 | 0.234 | 0.098 |
| HC vs. MDDr | 15.6 ± 3.0 | 15.5 ± 2.9 | -0.194 | 0.846 | 0.846 | -0.012 |
| MDDa vs. MDDr | 15.2 ± 3.2 | 15.5 ± 2.9 | -1421 | 0.156 | 0.234 | -0.111 |
| **Note.** The table shows the means and standard deviations of the mean number of streamlines within networks identified with Network-based statistics (NBS) toolbox from healthy control (HC) participants and patients with remitted (MDDr) or acute episode (MDDa) of Major depressive disorder when applying an NBS-threshold of F *=* 5.8 and F *=* 4.0. Test statistics are derived from pairwise t-tests and based on data contracted for sex, scanner-site, and age. CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance, HC = Healthy controls, MDDa = MDD patients with an acute episode, MDDr = MDD patients in symptomatic remission, T = Test statistic p = uncorrected p-value, pFDR = p-value corrected for multiple comparisons. |

|  |
| --- |
| **Table S11.T2***Partial correlations between depressive symptom severity, cognitive performance, and the sum of NOS within networks identified at NBS-threshold of F = 5.8 and F = 4.0* |
|  |  |  | **NBS *F-*threshold = 5.8** |  | **NBS *F-*threshold = 4.0** |
| **Network** |  |  | Sum of NOS | CF |  | Sum of NOS | CF |
| CF-PS network | Sum of NOS |  |   | 0.202\* |  |   | 0.219\* |
|  | HAMD |  | -0.105\* | -0.125\* |  | -0.081\* | -0.129\* |
| CF-VLM network | HC |  |   | -0.258\* |  |   | -0.251\* |
|  | MDDr |  | -0.035 | -0.107\* |  | -0.033 | -0.107\* |
| CF-VF network | HC |  |   | 0.048 |  |   | 0.048 |
|  | MDDr |  | -0.027 | -0.186\* |  | -0.042 | -0.185\* |
| **Note.** The table shows partial correlations between current depressive symptom severity (assessed by the Hamilton Depression Rating Scale (HAMD)), cognitive performance factor (CF), and the sum of NOS within networks identified with Network-based statistics (NBS) toolbox in patients with a remitted or acute episode of Major depressive disorder when applying an NBS-threshold of *F =* 5.8 and *F =* 4.0. Correlations were calculated as partial correlations between two of those variables while holding the third constant and correcting for age, sex, and study site. Significance was assessed through FDR Benjamini-Hochberg corrected p-values. All correlations that were considered significant at pFDR $\leq $.05 are marked with asterisks. CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance. |

# Supplement 12: Relationship between NBS F-threshold and network size

|  |
| --- |
| **Figure S12.F1***Size of the networks identified with NBS as a function of the F threshold applied in NBS* |
| C:\Users\m_grub03\owncloud\TraP\Projektordner\FactorAnalysis_Cognition_Connectome\3_Results\results\Fig_SizeNetworks_vs_Thresholds_noHM.jpg |
| **Note.** The figure shows the size of the networks identified with NBS when applying the respective F-threshold. NBS applies these thresholds to the test statistics representing the association between an edge and the respective CF. If an edge exceeds this threshold, it is included in the set of supra-threshold edges. The networks identified by NBS can only be composed of such supra-threshold edges. Networks for our analyses were extracted based on an F threshold of 5.8 (red line) because this threshold is the highest at which subnetworks could be identified for all cognitive factors, allowing both comparison of subnetworks and identification of the most specific subnetworks |

# Supplement 13: Results from exploratory mediation analysis

|  |
| --- |
| **Table S13.T1***Results from exploratory mediation analyses conducted to assess relationships between subnetwork-specific connectivity strength, current depressive symptom severity and cognitive performance* |
| **Path**  | **Coef.** | **SE** | **95%-CI LB** | **95%-CI UB** | **p-value** |
| *X = Total NOS in CF-PS network, M = Hamilton Depression Rating Scale, Y = CF-PS* |
| HAMD ~ X | -0.147 | 0.042 | -0.230 | -0.065 | 0.00 |
| Y ~ HAMD | -0.141 | 0.036 | -0.211 | -0.071 | 0.00 |
| Total | 0.223 | 0.039 | 0.146 | 0.299 | 0.00 |
| Direct | 0.206 | 0.039 | 0.129 | 0.282 | 0.00 |
| Indirect | 0.017 | 0.007 | 0.007 | 0.034 | 0.02 |
| *X = Total NOS in CF-VLM network, m = Hamilton Depression Rating Scale,Y = CF-VLM* |
| HAMD ~ X | -0.007 | 0.042 | -0.090 | 0.076 | 0.874 |
| Y ~ HAMD | -0.100 | 0.038 | -0.175 | -0.026 | 0.008 |
| Total | -0.269 | 0.040 | -0.349 | -0.190 | 0.000 |
| Direct | -0.270 | 0.040 | -0.349 | -0.191 | 0.000 |
| Indirect | 0.001 | 0.005 | -0.009 | 0.010 | 0.860 |
| *X = Total NOS in CF-VF network, M = Hamilton Depression Rating Scale, Y = CF-VF* |
| HAMD ~ X | -0.033 | 0.040 | -0.112 | 0.046 | 0.414 |
| Y ~ HAMD | -0.184 | 0.037 | -0.257 | -0.111 | 0.000 |
| Total | 0.059 | 0.039 | -0.018 | 0.136 | 0.135 |
| Direct | 0.053 | 0.039 | -0.023 | 0.129 | 0.172 |
| Indirect | 0.006 | 0.008 | -0.006 | 0.025 | 0.440 |
| **Note.** The table shows the results from mediation analyses conducted to assess the relationships between subnetwork-specific connectivity strength (NOS, independent variable), current depressive symptom severity as measured by the Hamilton depression rating scale (HAMD, mediator) and cognitive performance as measured by the cognitive factor (CF) used to identify the network. Path = Path of mediation analyses, Coef. = beta coefficient representing a given association, SE = standard error of the estimated coefficient, 95%-CI LB = Lower bound of the 95% confidence interval calculated for the coefficient, 95%-CI UB = Upper bound of the 95% confidence interval calculated for the coefficient, p-value = p-value representing the significance of the coefficient, i.e. its deviation of 0, CF-PS = CF representing processing speed performance, CF-VLM = CF representing verbal learning and memory performance, CF-VF = CF representing verbal fluency performance. |

# Supplement 14: Tests statistics representing between-group differences in cognitive performance

|  |
| --- |
| **Table S14.T1***Group differences in cognitive performance: Results from post-hoc t-tests* |
| **Cognitive Factor** | **HC (n = 821)** | **MDDr (n = 377)** | **MDDa (n = 315)** | **Test statistic** | **Significance1** |
| **CF-PS** | 0.244 ± 0.912 | -0.206 ± 0.99 | -0.388 ± 1.061 | 73.602 | A, B, C |
| **CF-VLM** | 0.127 ± 0.894 | -0.083 ± 0.991 | -0.228 ± 1.181 | 19.504 | A, B |
| **CF-VF** | 0.168 ± 0.963 | -0.054 ± 0.967 | -0.388 ± 0.974 | 38.188 | A, B, C |
| **Note.** The table shows the means and standard deviations of cognitive factors from healthy control (HC) participants and patients with remitted (MDDr) or acute episode (MDDa) of Major depressive disorder. Test statistics are derived from Analyses of Covariance controlling for sex and age. Significance was evaluated based on p-values from FDR (Benjamini-Hochberg) corrected post-hoc *t-*tests based on residualized data (corrected for age, sex, and years of education). CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance, HC = Healthy controls, MDDa = MDD patients with acute episode, MDDr = MDD patients in symptomatic remission, n.s. = not significant1 Letters indicate significant (i.e., pFDR < 0.05) differences between HC and MDDa (A), HC and MDDr (B), or MDDa and MDDr (C). |

|  |
| --- |
| **Table S14.T2***Effect sizes representing group differences in cognitive performance - sample grouped by remission status* |
|  |  |  | **Group 2** |
| **Cognitive factor** | **Group 1** |  | **MDDr** | **MDDa** |
| CF-PS | HC |  | 0.422 | 0.641 |
|  | MDDr |  |   | 0.198 |
| CF-VLM | HC |  | 0.154 | 0.305 |
|  | MDDr |  |   | 0.143 |
| CF-VF | HC |  | 0.239 | 0.568 |
|   | MDDr |   |   | 0.331 |
| **Note.** The table shows the effect sizes (Cohen’s d) representing differences in cognitive performance between healthy control (HC) participants and patients with remitted (MDDr) or acute episode (MDDa) of Major depressive disorder. CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance, HC = Healthy controls, MDDa = MDD patients with acute episode, MDDr = MDD patients in symptomatic remission. |

# Supplement 15: Differences in CF-VF when controlling for clinical covariates

|  |
| --- |
| **Table S15.T1***Differences between remission status subgroups in CF-VF when controlling for clinical covariates.* |
| **Covariate** | **MDDr** | **MDDa** | **T** | **p-value** | **Cohen’s d** |
| None / original results | 0.109 (0.902) | -0.132 (0.946) | 3.384 | 0.001 | 0.262 |
| Number of hospitalizations | 0.087 (0.890) | -0.106 (0.944) | 2.704 | 0.007 | 0.211 |
| Number of depressive episodes | 0.098 (0.905) | -0.121 (0.958) | 2.956 | 0.003 | 0.236 |
| Medication load | 0.090 (0.903) | -0.098 (0.936) | 2.164 | 0.031 | 0.204 |
| Age of disease onset | 0.103 (0.902) | -0.125 (0.947) | 3.168 | 0.002 | 0.246 |
| **Note.** The table shows means (standard deviations) of the cognitive factor representing verbal fluency performance (CF-VF) for MDD patients in symptomatic remission (MDDr) and acutely depressed MDD patients (MDDa) together with test statistics representing between-group differences. Test statistics were derived from post-hoc t-tests and based on data corrected for age, sex, years of education and the covariate depicted in the first column. |

# Supplement 16: Between-group differences in cognitive performance due to comorbidity

 To evaluate whether patients’ comorbidity had any significant influence on cognitive performance, we employed Analyses of Covariance within the subgroup of MDD patients, entering the cognitive factor score as dependent variable and comorbidity (yes/no) as independent variable while correcting for age, sex and years of education. We did not find any significant differences between patients with and without comorbid psychiatric disorders in any of the three cognitive factor scores.

|  |
| --- |
| **Table S16.T1***Group differences in cognitive performance due to comorbidity.* |
| **Cognitive factor** | **Comorbid (n = 293)** | **Non-Comorbid (n = 386)** | **Test statistic** | **p-value** |
| **CF-PS** | -0.025 ± 0.805 | 0.019 ± 0.945 | 0.408 | 0.523 |
| **CF-VLM** | 0.032 ± 0.962 | -0.025 ± 0.956 | 0.601 | 0.438 |
| **CF-VF** | -0.036 ± 0.919 | 0.028 ± 0.938 | 0.806 | 0.370 |
| **Note.** The table shows the means and standard deviations of cognitive factors from patients with and without comorbid psychiatric disorders. Means and standard deviations are corrected for influences of age, sex and years of education. Test statistics are derived from Analyses of Covariance controlling for sex, age, and years of education. CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance. |

# Supplement 17: Between-group differences in cognitive performance when excluding outliers

Since all cognitive factor (CF) included several outliers, we conducted several analyses to verify that between-group differences in cognitive performance were not only caused by those outliers. To this end, we applied outlier exclusion using the median absolute deviation z-score method (i.e. exclusion of all participants, whose absolute z-standardized value is > 3). Application of this method resulted in the exclusion of 13 participants (0.9%) from the analyses of CF-PS, 18 (1.2%) participants from the analysis of CF-VLM and 5 (0.3%) from the analysis of CF-VF. As shown in Table S17.T1, the significance of the differences between HCs and MDD patients remained mostly unchanged. The only exception is the difference between HC and acutely depressed individuals in CF-VLM which was significant in terms of uncorrected p-values (*p* = 0.020) but did not survive FDR correction (*p*FDR = 0.059).

|  |
| --- |
| **Table S17.T1***Group differences in cognitive performance previous to and after outlier exclusion: Results from post-hoc t-tests* |
| **Cognitive factor** | **HC** | **MDDr** | **MDDa** | **Test statistic** | **Significance1** |
| *Group differences previous to outlier exclusion* |
| **CF-PS** | 0.244 ± 0.912 | -0.207 ± 0.989 | -0.388 ± 1.059 | 79.093 | A, B |
| **CF-VLM** | 0.127 ± 0.899 | -0.084 ± 0.997 | -0.231 ± 1.19 | 20.61 | A |
| **CF-VF** | 0.172 ± 0.973 | -0.051 ± 0.977 | -0.39 ± 0.985 | 40.205 | A, B, C |
| *Group differences after excluding outliers* |
| **CF-PS** | 0.249 ± 0.888 | -0.19 ± 0.964 | -0.298 ± 0.917 | 70.561 | A, B |
| **CF-VLM** | 0.154 ± 0.843 | -0.047 ± 0.939 | -0.139 ± 1.06 | 17.238 | n.s. |
| **CF-VF** | 0.155 ± 0.943 | -0.051 ± 0.977 | -0.38 ± 0.97 | 37.788 | A, B, C |
| **Note.** The table shows the means and standard deviations of cognitive factors from healthy control (HC) participants and patients with remitted (MDDr) or acute episode (MDDa) of Major depressive disorder previous to and after outlier exclusion. Test statistics are derived from Analyses of Covariance controlling for sex, age, and years of education. Significance was evaluated based on p-values from FDR (Benjamini-Hochberg) corrected post-hoc *t-*tests based on residualized data (corrected for age, sex, and years of education). CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance, HC = Healthy controls, MDDa = MDD patients with acute episode, MDDr = MDD patients in symptomatic remission, n.s. = not significant1 Letters indicate significant (i.e., pFDR < 0.05) differences between HC and MDDa (A), HC and MDDr (B), or MDDa and MDDr (C). |

# Supplement 18: Remission status-related interaction effects

## Interaction effects on global-level cognition-connectome associations

|  |
| --- |
| **Table S18.T1.** *Remission status-related interaction effects for the association of global connectome measures and cognitive factor scores.* |
| **Connectome measure** | **Effect** | **F value** | **p-unc** | **pFDR** | **partial *η*²** |
| Total number of edges | Remission status x CF-VLM | 0.673 | 0.510 | 0.948 | 0.001 |
| Remission status x CF-PS | 0.077 | 0.926 | 0.948 | 0.000 |
| Remission status x CF-VF | 0.936 | 0.392 | 0.948 | 0.001 |
| Total number of streamlines | Remission status x CF-VLM | 0.081 | 0.922 | 0.948 | 0.000 |
| Remission status x CF-PS | 2.057 | 0.128 | 0.755 | 0.003 |
| Remission status x CF-VF | 0.485 | 0.616 | 0.948 | 0.001 |
| Mean number of streamlines per edge | Remission status x CF-VLM | 0.437 | 0.646 | 0.948 | 0.001 |
| Remission status x CF-PS | 2.738 | 0.065 | 0.755 | 0.004 |
| Remission status x CF-VF | 0.317 | 0.729 | 0.948 | 0.000 |
| Shortest path length | Remission status x CF-VLM | 1.754 | 0.173 | 0.755 | 0.002 |
| Remission status x CF-PS | 0.142 | 0.868 | 0.948 | 0.000 |
| Remission status x CF-VF | 0.307 | 0.736 | 0.948 | 0.000 |
| Global efficiency | Remission status x CF-VLM | 1.738 | 0.176 | 0.755 | 0.002 |
| Remission status x CF-PS | 0.125 | 0.882 | 0.948 | 0.000 |
| Remission status x CF-VF | 0.367 | 0.693 | 0.948 | 0.000 |
| Clustering coefficient | Remission status x CF-VLM | 0.981 | 0.375 | 0.948 | 0.001 |
| Remission status x CF-PS | 0.111 | 0.895 | 0.948 | 0.000 |
| Remission status x CF-VF | 0.602 | 0.548 | 0.948 | 0.001 |
| Normalized shortest path length | Remission status x CF-VLM | 2.306 | 0.100 | 0.755 | 0.003 |
| Remission status x CF-PS | 0.273 | 0.761 | 0.948 | 0.000 |
| Remission status x CF-VF | 0.061 | 0.941 | 0.948 | 0.000 |
| Normalized global efficiency | Remission status x CF-VLM | 2.457 | 0.086 | 0.755 | 0.003 |
| Remission status x CF-PS | 0.174 | 0.841 | 0.948 | 0.000 |
| Remission status x CF-VF | 0.053 | 0.948 | 0.948 | 0.000 |
| Normalized clustering coefficient | Remission status x CF-VLM | 1.743 | 0.175 | 0.755 | 0.002 |
| Remission status x CF-PS | 0.38 | 0.684 | 0.948 | 0.001 |
| Remission status x CF-VF | 0.562 | 0.570 | 0.948 | 0.001 |
| Small world index | Remission status x CF-VLM | 1.223 | 0.295 | 0.948 | 0.002 |
| Remission status x CF-PS | 0.785 | 0.456 | 0.948 | 0.001 |
| Remission status x CF-VF | 0.801 | 0.449 | 0.948 | 0.001 |
| ***Note.***The table shows the results from analyses testing the interaction effect of remission status and the cognitive factor scores on global connectome measures. P-unc = uncorrected p-values, pFDR = p-value corrected for multiple comparisons based on the False Discovery Rate. |

## Interaction effects on local-level cognition-connectome associations

|  |
| --- |
| **Table S18.T1.** *Results from NBS analyses on edge-wise interaction effects.* |
| **Network & Interaction** |  | **Number of edges** | ***p*FWE** |
| *CF-PS network* |  |  |  |
| Remission status x CF-PS |  | 46 | .070 |
|  |  |  |  |
| *CF-VLM network* |  |  |  |
| Remission status x CF-VLM |  | 89 | .076 |
|  |  |  |  |
| *CF-VF network* |  |  |  |
| Remission status x CF-VF |  | 27 | .149 |
| ***Note.*** The table shows the number of identified edges (i.e., the size of the largest network) associated to the remission status x cognitive performance interaction effect as well as the corresponding familywise error-corrected *p*-value. Note that the significance level was set to ɑ=.05/number of NBS analyses=.05/3 cognitive factors=.0167 in these analyses to correct for multiple testing. The analysis was conducted at an NBS *F*-threshold of *F*=5.8, which is in line with the NBS analyses of the association between local structural connectivity and cognitive performance. |

# Supplement 19: Description of identified networks

## Overlap of network edges

|  |
| --- |
| **Table S19.T1***Overlap of nodes and edges (in %) included in the subnetworks associated with cognitive factors.* |
| **Network** |  | **Overlap of network edges** | **Edges** | **Percent of the connectome** |
| **CF-PS** | **CF-VLM** | **CF-VF** |
| **CF-PS** |   |  | 10.39% | 16.88% | 77 | 1.20% |
| **CF-VLM** |  | 7.62% |  | 3.81% | 105 | 1.63% |
| **CF-VF** |   | 20.00% | 6.15% |   | 65 | 1.01% |
| **Note.** The table shows the overlap (in %) of edges included in the networks identified with Network-based statistics (NBS, *F-*threshold = 5.8). Overlap was calculated as number of edges of a given network A that are also part of a second network B divided by the total number of edges of A multiplied by 100. CF-PS = Network related to the cognitive factor (CF) representing processing speed performance, CF-VLM = Network related to the CF representing verbal learning and memory performance, CF-VF = Network related to the CF representing verbal fluency performance. Reading example: 10.39% of the 77 edges included in the CF-PS network are also included in the CF-VLM network. |

## Patterns of structural connectivity between brain regions

To further evaluate any differences in the patterns of structural connectivity within the identified subnetworks, we calculated the number of edges connecting frontal, temporal, parietal and occipital brain regions for each of the three subnetworks. As depicted in Table S19.T2 and in the heatmaps of Figure S19.F1, these numbers substantially differed across the subnetworks, indicating divergent patterns of structural connectivity.

|  |
| --- |
| **Table S19.T2***Number of edges from identified subnetworks connecting the brain regions.* |
|  |  |  | **Absolut number of edges** |  | **Percent of edges** |
| **Region 1** | **Region 2** |  | **CF-PS** | **CF-VLM** | **CF-VF** |  | **CF-PS** | **CF-VLM** | **CF-VF** |
| **frontal** | **frontal** |  | 20 | 17 | 6 |  | 25.97 | 16.19 | 9.23 |
|  | **temporal** |  | 7 | 6 | 11 |  | 9.09 | 5.71 | 16.92 |
|  | **parietal** |  | 12 | 17 | 9 |  | 15.58 | 16.19 | 13.85 |
|  | **occipital** |  | 3 | 2 | 6 |  | 3.90 | 1.90 | 9.23 |
| **temporal** | **temporal** |  | 5 | 4 | 10 |  | 6.49 | 3.81 | 15.38 |
|  | **parietal** |  | 7 | 26 | 11 |  | 9.09 | 24.76 | 16.92 |
|  | **occipital** |  | 12 | 11 | 6 |  | 15.58 | 10.48 | 9.23 |
| **parietal** | **parietal** |  | 5 | 12 | 3 |  | 6.49 | 11.43 | 4.62 |
|  | **occipital** |  | 3 | 8 | 0 |  | 3.90 | 7.62 | 0.00 |
| **occipital** | **occipital** |  | 3 | 2 | 3 |  | 3.90 | 1.90 | 4.62 |
| **Note.** The table shows the absolute and relative numbers of edges connecting the different brain regions across the subnetworks. |

|  |
| --- |
| **Figure S19.F1***Number of edges from identified networks connecting the brain regions.* |
| C:\Users\m_grub03\owncloud\TraP\Projektordner\FactorAnalysis_Cognition_Connectome\3_Results\results\Fig_Heatmaps_Connected_regions.png |
| **Note.** The figure shows the absolute numbers of edges connecting frontal, temporal, parietal and occipital brain regions across the identified subnetworks. CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance. |

## Involvement of left and right hemispheres

We conducted two complementary analyses on the lateralization of the effects. First, we compared the frequency of nodes from the left and right hemispheres participating in our networks. As can be seen from the table below, there were no significant differences regarding the involvement of nodes from the left and right hemispheres in any of the identified networks (ꭕ²(2) = 0.999, p = 0.607), indicating that nodes from both hemispheres contributed equally to the networks.

|  |
| --- |
| **Table S19.T3.** *Number (percent) of nodes from both hemispheres that participated in our networks.* |
|  | **Hemisphere** |  |
| **Network** | **Left** | **Right** | **All** |
| CF-PS | 32 (46.38%) | 37 (53.62%) | 69 |
| CF-VLM | 41 (51.90%) | 38 (48.10%) | 79 |
| CF-VF | 33 (55.00%) | 27 (45.00%) | 60 |
| **All** | 106 (50.96%) | 102 (49.04%) | 208 |
| ***Note.*** The table shows the number (proportion) of brain regions from left and right hemispheres that are included in the identified networks. CF-PS = network associated with processing speed factor, CF-VLM = network associated with verbal learning and memory factor, CF-VF = network associated with verbal fluency factor. |

Second, to evaluate the lateralization of the cognition-connectome association independent of the identified networks, we conducted an additional whole-brain analysis. To this end, we first calculated the mean connectivity strength of all 114 nodes included in our connectivity matrices as the mean number of streamlines across all edges that are connected to a given node. Second, for each of the three cognitive factor scores, we assigned each node a t-value representing its association with the respective factor score while correcting for age, sex, and scanner site. To evaluate any lateralization of the effects, we then compared 1) the distributions of the t-values assigned to the nodes from the left hemisphere to those assigned to nodes from the right hemisphere using Kolmogorov-Smirnov tests and 2) the mean effect size of the nodes from the left vs. right hemispheres by applying two-sample t-tests to the t-values. None of these tests yielded significant differences between the left and right hemispheres for any of our cognitive factor scores (see table below), again indicating no differences regarding the involvement of nodes from the left vs. right hemispheres.

|  |
| --- |
| **Table S19.T4.** *Results from whole-brain analyses on lateralization of cognition-connectome association.* |
|  | **Kolmogorov-Smirnov test****(distribution of t-values)** |  | **t-test****(effect size of t-values)** |
| **Factor** | **Test statistic** | ***p*-value** |  | **Test-statistic** | ***p*-value** |
| CF-PS | 0.123 | 0.788 |  | 0.082 | 0.935 |
| CF-VLM | 0.175 | 0.347 |  | 1.152 | 0.252 |
| CF-VF | 0.175 | 0.347 |  | 1.187 | 0.238 |
| ***Note.*** The table shows the results from our whole-brain analysis conducted to establish any lateralization regarding the cognition-connectome association. To this end, we aggregated the number of streamlines across all edges that are connected to one of the 114 nodes from our parcellation. For each cognitive factor, we then assigned each node a t-value representing its association with the respective factor score while correcting for age, sex, and scanner-site. The resulting t-values from nodes lying within the left or right hemispheres were compared with respect to their distribution (Kolmogorov-Smirnov test) and their mean effect size (t-test). |

In summary, neither the network-based nor the whole-brain approach did reveal significant differences regarding the involvement of nodes from the left and right hemispheres, indicating equal contributions of the left and right hemispheres to our cognitive factor scores.

# Supplement 20: Robustness checks of NBS analyses

To evaluate the impact of participants’ years of education and head motion during MRI acquisition on the identified networks, we used linear regression models to test for any significant associations between subnetwork-specific connectivity strength and one of these variables while correcting for age, sex, scanner-site and the cognitive factor score used to identify the respective subnetwork. We obtained head motion from FSL's eddy\_movement\_rms output by averaging the volume-specific displacement (defined as the square root of the average squared displacement of each voxel within a given volume relative to the previous volume) across all volumes from a given participant. Neither participants’ years of education nor head motion was significantly associated with any of the measures for subnetwork-specific connectivity strength (see Table S20.T1).

|  |
| --- |
| **Table S20.T1***Associations between connectivity strength in identified networks and years of education or head motion.* |
|  |  | **Years of education** |  | **Head motion** |
| **Network** |  | **Beta** | **p** |  | **Beta** | **p** |
| CP-PS |  | -0.011 | 0.648 |  | -0.064 | 0.162 |
| CP-VLM |  | 0.027 | 0.277 |  | -0.005 | 0.914 |
| CP-hEF |  | 0.030 | 0.256 |  | -0.022 | 0.654 |
| **Note.** The table shows associations between subnetwork-specific connectivity strength and participants’ years of education or head motion during MRI acquisition. Beta = standardized regression coefficients, p = uncorrected p-value, CF-PS = Network related to the cognitive factor representing processing speed performance, CF-VLM = Network related to the cognitive factor representing verbal learning and memory performance, CF-VF = Network related to the Cognitive factor representing verbal fluency performance. |

To assess the influence of outliers within the cognitive factor scores and our decision of using thresholded connectivity matrices, we rerun our NBS analyses 1) after excluding outliers and 2) based on non-thresholded connectivity matrices. Participants were considered outliers, if their absolute z-standardized value of the cognitive factor was > 3. None of these approaches changed the overall pattern of our results, as shown in Table S20.T2.

|  |
| --- |
| **Table S20.T2***Associations captured within the subnetworks when excluding outliers or using non-thresholded connectivity matrices.* |
| **Network** | **Analyses based on** | **t** | **p** | **η²** |
| CF-PS | Outlier exclusion | 6.493 | < 0.001 | 0.0280 |
|  | Non-threshold Connectivity matrices | 8.548 | < 0.001 | 0.0471 |
| CF-VLM | Outlier exclusion | -9.124 | < 0.001 | 0.0539 |
|  | Non-threshold Connectivity matrices | -9.592 | < 0.001 | 0.0586 |
| CF-VF | Outlier exclusion | 3.650 | < 0.001 | 0.0086 |
|  | Non-threshold Connectivity matrices | 2.270 | 0.023 | 0.0035 |
| **Note.** The table shows the association between subnetwork-specific connectivity strength (i.e. mean number of streamlines within edges included in a given subnetwork) and the cognitive factor used to identify the subnetwork. T = t-value representing the significance of the association, p = uncorrected p-values, η² = partial eta squared, CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance. |

# Supplement 21: Full results from correlational analyses

|  |
| --- |
| **Table S21.T1***Correlations between cognitive factors, subnetwork-specific connectivity strength and depressive symptoms* |
| **Network** |  | Sum of NOS | CF |
| CF-PS | Sum of NOS |   | 0.201\* |
|  | HAMD | -0.105\* | -0.126\* |
| CF-VLM | Sum of NOS |   | -0.251\* |
|  | HAMD | -0.033 | -0.107\* |
| CF-VF | Sum of NOS |   | 0.053 |
|   | HAMD  | -0.021 | -0.186\* |
| **Note.** The table shows the results from correlational analyses conducted to establish the relationships between connectivity strength within identified subnetworks (Sum of NOS within the networks identified when using an NBS *F-*threshold of *F =* 5.8), current depressive symptom severity (Hamilton Depression Rating Scale, HAMD) and cognitive performance as measured by the cognitive factors (CF) used to identify the subnetwork. Correlations were calculated as partial correlations between two variables while correcting for the third variable, age, sex, study site.$$\* =p\_{FDR}\leq 0.05$$ |

# Supplement 22: Symptom specific partial correlations

|  |
| --- |
| **Table S22.T1.** *Symptom-specific partial correlations with connectivity in processing speed network.* |
| **X** | **Y** | **r** | **CI95%** | **punc** | **pFDR** |
| Connectivity strength in processing speed network | Depressed mood | -0.039 | [-0.11 0.04] | 0.312 | 0.468 |
| Work and activities | -0.055 | [-0.13 0.02] | 0.151 | 0.352 |
| Genital symptoms | -0.087 | [-0.16 -0.01] | 0.024 | 0.126 |
| Somatic symptoms - gastrointestinal |  0.019 | [-0.06 0.09] | 0.622 | 0.725 |
| Loss of weight | -0.059 | [-0.13 0.02] | 0.129 | 0.339 |
| Sleep onset insomnia | -0.051 | [-0.13 0.02] | 0.189 | 0.360 |
| Sleep maintenance insomnia | -0.029 | [-0.10 0.05] | 0.446 | 0.552 |
| Early awakening | -0.008 | [-0.08 0.07] | 0.841 | 0.883 |
| Somatic symptoms general | -0.044 | [-0.12 0.03] | 0.254 | 0.411 |
| Feelings of guilt | -0.097 | [-0.17 -0.02] | 0.012 | 0.081 |
| Suicide | -0.059 | [-0.13 0.02] | 0.124 | 0.339 |
| Anxiety Psychic | -0.107 | [-0.18 -0.03] | 0.005 | 0.081 |
| Anxiety somatic | -0.033 | [-0.11 0.04] | 0.392 | 0.549 |
| Hypochrondriasis | -0.014 | [-0.09 0.06] | 0.724 | 0.800 |
| Insight |  0.045 | [-0.03 0.12] | 0.239 | 0.411 |
| Retardation | -0.053 | [-0.13 0.02] | 0.170 | 0.358 |
| Agitation | -0.063 | [-0.14 0.01] | 0.105 | 0.339 |
| Diurnal variation Type A | -0.068 | [-0.14 0.01] | 0.078 | 0.328 |
| Depersonalization and derealization |  0.030 | [-0.05 0.11] | 0.440 | 0.552 |
| Paranoid symptoms | -0.101 | [-0.18 -0.03] | 0.009 | 0.081 |
| Obsessional and compulsive symptoms | -0.004 | [-0.08 0.07] | 0.914 | 0.914 |
| ***Note.*** The table shows the correlations between each symptom of the 21-item Hamilton Depression Rating Scale and the connectivity strength within the processing speed network. Correlations were calculated as Spearman partial correlations between a given symptom and the connectivity strength while correcting for age, sex, scanner-site, and the CF-PS performance. r = Spearman partial correlation coefficient, CI95% = 95% parametric confidence interval of the correlation coefficient, punx = uncorrected p-value, pFDR = p-value corrected for multiple comparisons based on the FDR. |

# Supplement 23: Correlational analyses correcting for clinical covariates

|  |
| --- |
| **Table S22.T1***Correlations between depressive symptoms and clinical variables used as covariates in the following analysis.* |
| **X** | **Y** | **r** | **95%-CI** | **p** | **PFDR** |
| HAMD | Medication load index | 0.390 | [0.31, 0.47] | 0.000 | 0.000 |
| Number of depressive episodes | 0.160 | [0.08, 0.24] | 0.000 | 0.000 |
| Number of hospitalizations | 0.281 | [0.21, 0.35] | 0.000 | 0.000 |
| Age of disease onset | -0.050 | [-0.13, 0.03] | 0.198 | 0.198 |
| **Note.** The table shows correlations between current depressive symptom severity (Hamilton Depression Rating Scale, HAMD) and the medication load, the number of depressive episodes, number of hospitalizations or age of disease onset. Correlations were calculated as partial correlations between two variables while correcting for age and sex. To estimate medication load, we calculated an established Medication Load Index (48–50), which reflects dose and number of prescriptions irrespective of active components. r = correlation coefficient, 95%-CI = 95% confidence interval of the correlation coefficient, p = uncorrected p value, pFDR = p-value corrected for multiple comparisons using the FDR Benjamini-Hochberg procedure. |

|  |
| --- |
| **Table S22.T2** *Results from correlational analyses when controlling for clinical covariates* |
| **Network** |  |  | **Original results** |  | **Corrected for Medication load index** |  | **Corrected for number of depressive episodes** |  | **Corrected for number of hospitalizations** |  | **Corrected for age of disease onset** |
|  |  | Sum of NOS | CF |   | Sum of NOS | CF |  | Sum of NOS | CF |   | Sum of NOS | CF |  | Sum of NOS | CF |
| CF-PS | Sum of NOS |  |   | 0.201\* |  |   | 0.179\* |  |   | 0.192\* |  |   | 0.202\* |  |   | 0.202\* |
|  | HAMD |  | -0.105\* | -0.126\* |  | -0.1 | -0.065 |  | -0.104\* | -0.104\* |  | -0.105\* | -0.086\* |  | -0.105\* | -0.125\* |
| CF-VLM | Sum of NOS |  |   | -0.251\* |  |   | -0.26\* |  |   | -0.243\* |  |   | -0.252\* |  |   | -0.258\* |
|  | HAMD |  | -0.033 | -0.107\* |  | -0.013 | -0.023 |  | -0.026 | -0.1\* |  | -0.033 | -0.073 |  | -0.035 | -0.107\* |
| CF-VF | Sum of NOS |  |   | 0.053 |  |   | 0.062 |  |   | 0.053 |  |   | 0.05 |  |   | 0.048 |
|   | HAMD  |  | -0.021 | -0.186\* |  | -0.029 | -0.126\* |  | -0.021 | -0.171\* |  | -0.013 | -0.145\* |  | -0.027 | -0.186\* |
| **Note.** The table shows the results from correlational analyses conducted to establish the relationships between connectivity strength within identified subnetworks (Sum of NOS within the networks identified when using an NBS *F-*threshold of *F =* 5.8), current depressive symptom severity (Hamilton Depression Rating Scale, HAMD) and cognitive performance as measured by the cognitive factors (CF) used to identify the subnetwork. Correlations were calculated as partial correlations between two variables while correcting for the third variable, age, sex, study site and either the medication load, the number of depressive episodes, number of hospitalizations or age of disease onset. To estimate medication load, we calculated an established Medication Load Index (48–50), which reflects dose and number of prescriptions irrespective of active components. P-values representing the significance of the correlations were extracted and corrected for multiple comparisons using the FDR Benjamini-Hochberg procedure. CF-PS = Cognitive factor representing processing speed performance, CF-VLM = Cognitive factor representing verbal learning and memory performance, CF-VF = Cognitive factor representing verbal fluency performance.\* $ =p\_{FDR}\leq 0.05$ |

# Supplementary references

1. Helmstaedter C, Lendt M, Lux S (2001): Verbaler Lern-und Merkfähigkeitstest: VLMT; Manual. *Beltz-Test*.

2. Jaeger J (2018): Digit symbol substitution test. In: Kreutzer J, Caplan B, DeLuca J, editors. *Journal of Clinical Psychopharmacology*, vol. 38. New York, London: Springer, pp 513–519.

3. Wechsler D (2008): Wechsler Adult Intelligence Scale - Fourth Edition: Administration and Scoring Manual. *Psychological Corporation*. New York: Oxford University Press.

4. Reitan R (1956): Trail Making test: Manual for administration, scoring, and interpretation. *Bloomington: Indiana University*. Indianapolis: Indiana University Press.

5. Ross RM (2005): *The D2 Test of Attention: An Examination of Age, Gender, and Cross-Cultural Indices*. Argosy University. Retrieved from https://books.google.dk/books/about/The\_D2\_Test\_of\_Attention.html?id=yEz-MQAACAAJ&redir\_esc=y

6. Kessels RPC, Van Den Berg E, Ruis C, Brands AMA (2008): The backward span of the corsi block-tapping task and its association with the WAIS-III digit span. *Assessment* 15: 426–434.

7. Aschenbrenner S, Tucha O, Lange K (2000): Regensburger Wortflüssigkeitstest. *Göttingen: Hogrefe*.

8. Andersson JLR, Skare S (2002): A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI. *Neuroimage* 16: 177–199.

9. Andersson JLR, Sotiropoulos SN (2016): An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125: 1063–1078.

10. Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, *et al.* (2009): Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45. https://doi.org/10.1016/j.neuroimage.2008.10.055

11. de Lange SC, Heuvel MP van den (2021): Structural and functional connectivity reconstruction with CATO - A Connectivity Analysis TOolbox. *bioRxiv* 2021.05.31.446012.

12. Chang LC, Jones DK, Pierpaoli C (2005): RESTORE: Robust estimation of tensors by outlier rejection. *Magn Reson Med* 53: 1088–1095.

13. Chang LC, Walker L, Pierpaoli C (2012): Informed RESTORE: A method for robust estimation of diffusion tensor from low redundancy datasets in the presence of physiological noise artifacts. *Magn Reson Med* 68: 1654–1663.

14. Van Den Heuvel MP, Sporns O, Collin G, Scheewe T, Mandl RCW, Cahn W, *et al.* (2013): Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA Psychiatry* 70: 783–792.

15. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Van Wedeen J, Sporns O (2008): Mapping the structural core of human cerebral cortex. *PLoS Biol* 6: 1479–1493.

16. Cammoun L, Gigandet X, Meskaldji D, Thiran JP, Sporns O, Do KQ, *et al.* (2012): Mapping the human connectome at multiple scales with diffusion spectrum MRI. *J Neurosci Methods* 203: 386–397.

17. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, *et al.* (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31: 968–980.

18. Repple J, Mauritz M, Meinert S, de Lange SC, Grotegerd D, Opel N, *et al.* (2020): Severity of current depression and remission status are associated with structural connectome alterations in major depressive disorder. *Mol Psychiatry* 25: 1550–1558.

19. de Lange SC, Scholtens LH, van den Berg LH, Boks MP, Bozzali M, Cahn W, *et al.* (2019): Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. *Nat Hum Behav* 3: 988–998.

20. Mori S, van Zijl P (2002): Fiber tracking: principles and strategies - a technical review. *NMR Biomed* 15: 468–480.

21. Sarwar T, Ramamohanarao K, Zalesky A (2019): Mapping connectomes with diffusion MRI: deterministic or probabilistic tractography? *Magn Reson Med* 81: 1368–1384.

22. Zalesky A, Fornito A, Cocchi L, Gollo LL, van den Heuvel MP, Breakspear M (2016): Connectome sensitivity or specificity: which is more important? *Neuroimage* 142: 407–420.

23. de Reus MA, van den Heuvel MP (2013): Estimating false positives and negatives in brain networks. *Neuroimage* 70: 402–409.

24. Maldonado T, Orr JM, Goen JRM, Bernard JA (2020): Age differences in the subcomponents of executive functioning. *Journals Gerontol - Ser B Psychol Sci Soc Sci* 75: e31–e55.

25. Lipnicki DM, Crawford JD, Dutta R, Thalamuthu A, Kochan NA, Andrews G, *et al.* (2017): Age-related cognitive decline and associations with sex, education and apolipoprotein E genotype across ethnocultural groups and geographic regions: a collaborative cohort study. *PLoS Med* 14. https://doi.org/10.1371/journal.pmed.1002261

26. Rhodes S, Greene NR, Naveh-Benjamin M (2019): Age-related differences in recall and recognition: a meta-analysis. *Psychonomic Bulletin and Review*, vol. 26. pp 1529–1547.

27. Gong G, Rosa-Neto P, Carbonell F, Chen ZJ, He Y, Evans AC (2009): Age- and gender-related differences in the cortical anatomical network. *J Neurosci* 29: 15684–15693.

28. Robinson EC, Hammers A, Ericsson A, Edwards AD, Rueckert D (2010): Identifying population differences in whole-brain structural networks: A machine learning approach. *Neuroimage* 50: 910–919.

29. Zhao T, Cao M, Niu H, Zuo XN, Evans A, He Y, *et al.* (2015): Age-related changes in the topological organization of the white matter structural connectome across the human lifespan. *Hum Brain Mapp* 36: 3777–3792.

30. Buchanan CR, Bastin ME, Ritchie SJ, Liewald DC, Madole JW, Tucker-Drob EM, *et al.* (2020): The effect of network thresholding and weighting on structural brain networks in the UK Biobank. *Neuroimage* 211: 116443.

31. Van Den Heuvel MP, Scholtens LH, Van Der Burgh HK, Agosta F, Alloza C, Arango C, *et al.* (2019): 10kin1day: A bottom-up neuroimaging initiative. *Front Neurol* 10: 425.

32. Rubinov M, Sporns O (2010): Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* 52: 1059–1069.

33. Biggs J, Madnani N (2017): factor\_analyzer package [no. 0.3.1]. Retrieved from https://factor-analyzer.readthedocs.io/en/latest/factor\_analyzer.html

34. Hunter JD (2007): Matplotlib: A 2D graphics environment. *Comput Sci Eng* 9: 90–95.

35. Van Der Walt S, Colbert SC, Varoquaux G (2011): The NumPy array: A structure for efficient numerical computation. *Comput Sci Eng* 13: 22–30.

36. McKinney W (2010): Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*. https://doi.org/10.25080/majora-92bf1922-00a

37. Vallat R (2018): Pingouin: statistics in Python. *J Open Source Softw* 3: 1026.

38. Welcome to researchpy’s documentation! — researchpy 0.3.2 documentation (n.d.): Retrieved July 26, 2021, from https://researchpy.readthedocs.io/en/latest/

39. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, *et al.* (2020): SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 17: 261–272.

40. Waskom M (2021): Seaborn: Statistical Data Visualization. *J Open Source Softw* 6: 3021.

41. Seabold S, Perktold J (2010): Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*. https://doi.org/10.25080/majora-92bf1922-011

42. Moosbrugger H, Schermelleh-Engel K (2012): Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). 325–343.

43. Horn JL (1965): A rationale and test for the number of factors in factor analysis. *Psychometrika* 30: 179–185.

44. Team R Development Core (2018): A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, vol. 2. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org

45. W R (2021): *psychTools:Tools to Accompany the ’Psych; Package for Psychological Research*. Evanston, Illinois: Northwestern University.

46. Zalesky A, Fornito A, Bullmore ET (2010): Network-based statistic: Identifying differences in brain networks. *Neuroimage* 53: 1197–1207.

47. Freedman D, Lane D (1983): A nonstochastic interpretation of reported significance levels. *J Bus Econ Stat* 1: 292–298.

48. Opel N, Redlich R, Dohm K, Zaremba D, Goltermann J, Repple J, *et al.* (2019): Mediation of the influence of childhood maltreatment on depression relapse by cortical structure: a 2-year longitudinal observational study. *The Lancet Psychiatry* 6: 318–326.

49. Redlich R, Almeida JR, Grotegerd D, Opel N, Kugel H, Heindel W, *et al.* (2014): Brain morphometric biomarkers distinguishing unipolar and bipolar depression: A voxel-based morphometry-pattern classification approach. *JAMA Psychiatry* 71: 1222–1230.

50. Repple J, Meinert S, Grotegerd D, Kugel H, Redlich R, Dohm K, *et al.* (2017): A voxel-based diffusion tensor imaging study in unipolar and bipolar depression. *Bipolar Disord* 19: 23–31.

1. We chose this threshold because it is the highest at which subnetworks could be identified for all cognitive factors, allowing both comparison of subnetworks and identification of the most specific subnetworks. [↑](#footnote-ref-1)