**Supporting Information**

**Supporting Methods**

**Mediation analysis**

TFCE\_mediation both incorporate cortex-wise mediation analysis with TFCE enhancement. Mediation models are statistically causal models that can determine the effect of an independent variable on two dependent variables ([Baron and Kenny, 1986](#_ENREF_2)). In the case of complete mediation, the effect of the independent variable on the dependent variable is explained entirely by a mediator variable. This variance explained by a mediator variable is called the indirect effect. In practice, complete mediation is rare. Alternatively, partial mediation occurs when only some of the variance is explained. Cortex-wise mediation applies a simple model of mediation among each voxel or vertex, and examines if there is a significant indirect effect. Importantly, cortex-wise mediation also provides spatial information, as there must be a relationship among each variable for a significant indirect effect to occur.

The standard assumptions for traditional mediation should be applied also to cortex-wise mediation ([Baron and Kenny, 1986](#_ENREF_2)). Prior to the mediation analysis, the independent variable should be significantly associated with the mediator variable and the dependent variable, as well as the mediator variable should be significantly associated with the dependent variable with the independent variable as a covariate. Therefore, in any analysis, two significant whole brain corrected associations (PFWE<0.05) should exist prior to cortex-wise mediation. In the best case scenario, there is an overlap among the cortex-wise measures associated in each of the two significant whole brain associations. However, given the stringent multiple comparison correction for whole brain analyses, significant cortex-wise mediation may occur even in locations that were not significant independently.

The mediation model also depends heavily on strong *a priori* hypotheses. For statistical causal inference, the independent variable should be truly independent from the mediator and dependent variables. Moreover, the mediator should be independent from the dependent variable. The independence cannot be determined statistically, rather it must be defined. For example, genotype as an independent variable predicts neurocognitive performance as the dependent variable (Figure 1; Path C). Furthermore, genotype predicts the mediator voxel-wise FA (Path A), as well as voxel-wise FA predicts cognitive performance (Path B). In this case, a clear direction of the indirect effect (via path A and B) can be asserted. If the indirect effect is significant, it suggests that the effect of genotype on neurocognitive performance is mediated by the subset of voxels associated with cognitive performance. Importantly, as already stated, mediation would only occur at voxels in which genotype predicts FA and FA predicts cognitive performance.

Given that cortex-wise mediation provides spatial information of an association, it may be used beyond statistical causal inference. It tests for a common association among different neuroimaging modalities. For instance, it could be determined which FA voxels are mediating the effect of genotype on dorsolateral prefrontal cortex (DLPFC) surface area. If significant voxel-wise mediation occurs, it could be inferred that genotype affects these FA voxels and these voxels also predict DLPFC surface area. That is, it can be used to integrate associations among different neuroimaging phenotypes.

There are three models of cortex-wise mediation that can be tested using TFCE\_mediation: (1), cortex-wise image as the independent variable; (2), cortex-wise image as the mediator variable (as in Figure 1), and (3) cortex-wise image as the dependent variable. Two sets of regression are performed to assess the indirect effect using the Aroian variant of the Sobel equation ([Mackinnon, et al., 1995](#_ENREF_17); [Sobel, 1982](#_ENREF_25); [Sobel, 1986](#_ENREF_26)). The regression analyses are performed using the same two-step regression analysis previously described. The default version is the Aroian variant because it does not assume the product of Sa and Sb is small.

The unstandardized regression coefficients (betas: a and b) and the standard errors (Sa and Sb) from the independent variable are regressed on the mediator variable (Path A), and the mediator variable to the dependent variable (Path B) to produce a z-value at each cortex-wise measure (Figure S1). TFCE\_mediation usesan equivalent Sobel equation using t-values (ta and tb) in order to vectorize the regression analyses.

**Image Acquisition**

In the MooDS sample, structural T1-weighted MRI scans were acquired on three Siemens Trio 3T MR (Siemens, Erlangen, Germany) scanners at Charité - Universitätsmedizin Berlin, the Life and Brain Center of the University of Bonn, and the Central Institute of Mental Health Mannheim. Structural scans were acquired using a T1-weighted three-dimensional magnetization prepared rapid gradient echo (MP-RAGE) sequence with an isotropic spatial resolution of 1 mm3 (repetition time (TR) = 1.57s, echo time (TE) = 2.74ms, flip angle = 15°). Quality assurance (QA) measurements were conducted at all three study sites on every day of data collection utilizing a multicenter quality assurance protocol (Friedman and Glover, 2006), which revealed stable signals over time and comparable quality between sites. In the NGFN\_PLUS sample, structural scans were also acquired using a T1-weighted three-dimensional MP-RAGE sequence with an isotropic spatial resolution of 1 mm3. At Charité Berlin, ten subjects were scanned using a Siemens Trio 3T MR and 45 subjects were scanned using a MAGNETOM Verio 3T (Siemens, Erlangen, Germany) using the same acquisition parameters (TR=2.3s, TE=3.03ms, flip angle = 9°). At the Life and Brain Center of the University of Bonn, subjects were scanned using a Siemens Trio 3T MR (TR=2.3s, TE=3.93ms, flip angle = 9°). At the Central Institute of Mental Health Mannheim, subjects were scanned using a Siemens Trio 3T MR (TR = 2.3s, TE = 3.03ms, flip angle = 9°).

BOLD fMRI was also performed on both samples. Identical sequences and scanner protocols were used across sites within each sample (MooDS: 28 slices, 4 mm/slices, 1 mm gap, FOV 192 mm, matrix = 64 × 64, TR=2s, TE=30ms, flip angle 90°; NGFN\_PLUS: 28 slices, 4 mm/slices, 1 mm gap, FOV 192 mm, matrix = 64 × 64, TR=2s, TE=30ms, flip angle 80°). QA measures were conducted on every measurement day at both sites according to a multicenter QA protocol revealing stable signals over time, and described in detail in previous publications ([Esslinger, et al., 2009](#_ENREF_8)). During functional MRI scanning, subjects completed the N-Back task (see below).

In the MooDS sample, diffusion-weighted MRI (DW-MRI) image acquisition was additionally performed. DW-MRI images collected at Charité - Universitätsmedizin Berlin had 2mm isotropic voxel dimensions, 61 slices, 64 gradient non-collinear directions with a b value = 1000 s/mm2, TR =8200 ms, TE = 92 ms. DW-MRI collected at the Life and Brain Center of the University of Bonn had a spatial resolution of 1.7mm isotropic, 72 slices, 60 non-collinear gradient directions with a b value = 1000 s/mm2, TR =12000 ms, TE = 100 ms. DW-MRI images collected at the Central Institute of Mental Health Mannheim had a spatial resolution of 2mm isotropic, 64 slices, 60 gradient, non-collinear directions with a b value = 1000 s/ mm2, TR =14000 ms, TE = 86 ms. DW-MRI was not performed in the NGFN\_PLUS sample.

**Structural Image Processing**

Cortical reconstruction was performed on all T1-weighted images using the Freesurfer image analysis suite (http://surfer.nmr.mgh.harvard.edu/). The technical details of these procedures are described in prior publications ([Fischl and Dale, 2000](#_ENREF_9); [Fischl, et al., 2002](#_ENREF_10); [Fischl, et al., 2004](#_ENREF_11); [Fischl, et al., 1999a](#_ENREF_12); [Fischl, et al., 1999b](#_ENREF_13)). In brief, this process includes removal of non-brain tissue, segmentation of the subcortical white matter and deep gray matter volumetric structures, intensity normalization, tessellation of the gray matter white matter boundary, automated topology correction, and surface deformation. A number of deformable procedures were performed including surface inflation, registration to a spherical atlas which is based on individual cortical folding patterns to match cortical geometry across subjects, and creation of a variety of surface based data including maps of curvature and sulcal depth. Both intensity and continuity information from the entire three dimensional MR volume in segmentation and deformation procedures to produce representations of cortical thickness, calculated as the closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the tessellated surface ([Fischl and Dale, 2000](#_ENREF_9)).

**Functional Image Processing**

**N-Back Task**

Details of the N-back protocol have been published elsewhere ([Charlet, et al., 2014](#_ENREF_6); [Esslinger, et al., 2009](#_ENREF_8)). Briefly, a visuospatial version of the N-back task (0/2-back version) was used to probe the working memory network. The N-back robustly engages the DLPFC and inferior parietal lobule. It uses a block design with four 30 seconds blocks of the control (0-back) condition with low working memory load and four 30 second blocks of the cognitively high demand- working memory (2-back) condition; each consisting of 14 trials (sequential presentation of a digit (1-4) for 500ms with a interstimulus interval of 1500ms). During the control condition, subjects were asked to respond on their response box (Fiber Optic Computer Response Devices for fMRI research, http://www.curdes.com/) correspondingly to the digit displayed. In the 2-back condition they had to respond correspondingly to the digit seen two trials earlier while continuously encoding sequentially displayed digits.

**fMRI Analysis**

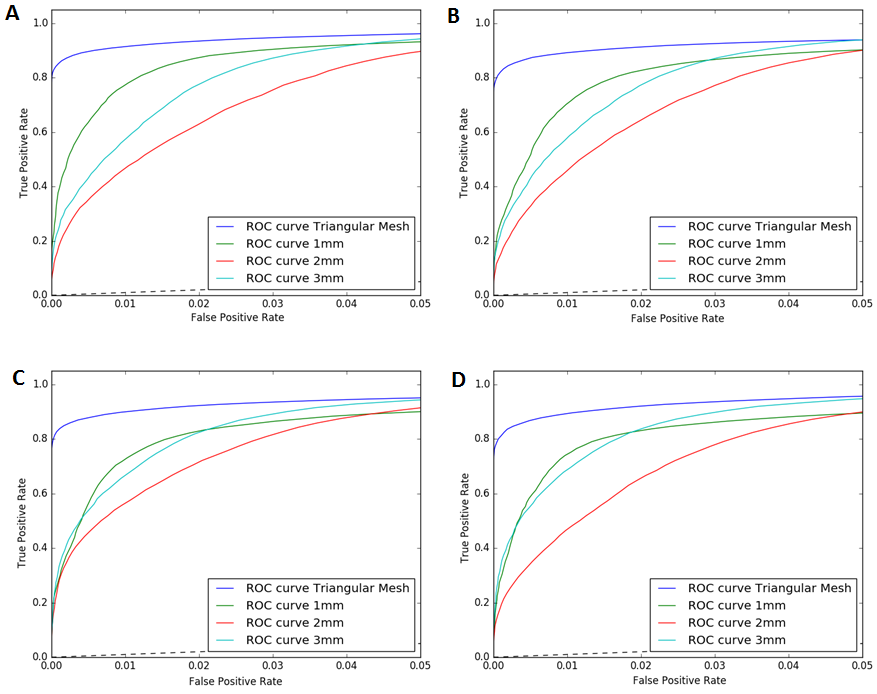
The fMRI analysis was the same across all samples and sites. Preprocessing was performed using FSL FEAT version 6.0.0 ([Smith, et al., 2004](#_ENREF_24)). Each 4D image of 125 volumes was motion corrected, adjusted for slice timing differences, and spatially smoothed (8mm FWHM). Next, functional data were registered to the high resolution T1 image (using boundary based registration in FSL FLIRT) ([Greve and Fischl, 2009](#_ENREF_14); [Jenkinson, et al., 2002](#_ENREF_16)), and the high resolution T1 image underwent a non-linear transformation to the MNI152\_T1\_2mm template (FSL FNIRT). Automatic artefact correction based on independent component analysis on the individual 4D images (FSL Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) version 3.0) was performed using ICA-AROMA version 0.3B ([Beckmann and Smith, 2004](#_ENREF_3); [Pruim, et al., 2015a](#_ENREF_20); [Pruim, et al., 2015b](#_ENREF_21)). Last, each 4D image was high pass filtered (100s).

To provide a summary measure of coherent task activation, we performed group tensor ICA using FSL MELODIC ([Beckmann and Smith, 2005](#_ENREF_4); [Guo and Pagnoni, 2008](#_ENREF_15)). All images were down-sampled to 4mm isotropic voxels to reduce computational burden. Group independent components were extracted across the space, time, and subject dimensions, testing against the N-back design and two main task contrasts. From the resulting components, we chose the component that best fit the expected 2-back greater than 0-back spatial activation ([Owen, et al., 2005](#_ENREF_19)), and we used the subject loadings on this component as the general measure of how well each subject fits the pattern of coherent task activation ([Beckmann and Smith, 2005](#_ENREF_4); [Calhoun, et al., 2009](#_ENREF_5)).

**DW-MRI Processing**

After visual inspection, all DW-MRI scans underwent automated quality control, eddy current, and motion correction using DTIprep (<https://www.nitrc.org/projects/dtiprep/>) ([Oguz, et al., 2014](#_ENREF_18)). Since DW-MRI scans were collected at 3T, we further performed a correction for EPI-induced susceptibility artifacts using high dimension warping of average b0 images to a T1-weighted scan for the same subject. T1-weighted scans underwent inhomogeneity correction with ANTS’s N4BiasFieldCorrection ([Tustison, et al., 2010](#_ENREF_27)), were cropped using FSL robustfov and skull stripped using BET with robust brain center estimation, and neck cleanup ([Smith, 2002](#_ENREF_22)). The b0 images were averaged using mcflirt, and also skull stripped using BET. Using FSL flirt, the T1-weighted images were then linear transformed to MNI152\_T1\_1mm\_brain standard image, and the b0 images were linear transformed to the T1-weight images using no shearing parameters ([Greve and Fischl, 2009](#_ENREF_14)). Using ANTS, a non-linear transformation was performed between the images, and the resulting field map transformation was applied to the DW-DTI image ([Avants, et al., 2008](#_ENREF_1)). Each EPI distortion corrected image was visually inspected. Fractional anisotropy (FA) images were created by fitting a tensor model at each voxel using FSL DTIFit ([Smith, et al., 2004](#_ENREF_24)). Voxel-wise analysis of the FA data was carried out using TBSS version 1.2 ([Smith, et al., 2006](#_ENREF_23)). FA images first underwent nonlinear registration to the FMRIB58\_FA target image. The mean FA image was iteratively generated. Each group was then aligned to MNI 152 standard space using an affine transformation. An average white matter skeleton was then generated from the mean of all subjects’-transformed FA images at a threshold of 0.2, and each subject’s FA data was projected onto the white matter skeleton. For comparison to TFCE\_mediation, voxel-wise statistics were calculated using randomise (v2.1) with 10000 permutations ([Winkler, et al., 2014](#_ENREF_28)). To directly compare the multiple regression results from FSL randomise and TFCE\_mediation, Dice similarity coefficients among the results were assessed using binarized masks of three thresholds: PFWE<0.1, PFWE<0.05, PFWE<0.01 ([Dice, 1945](#_ENREF_7)).

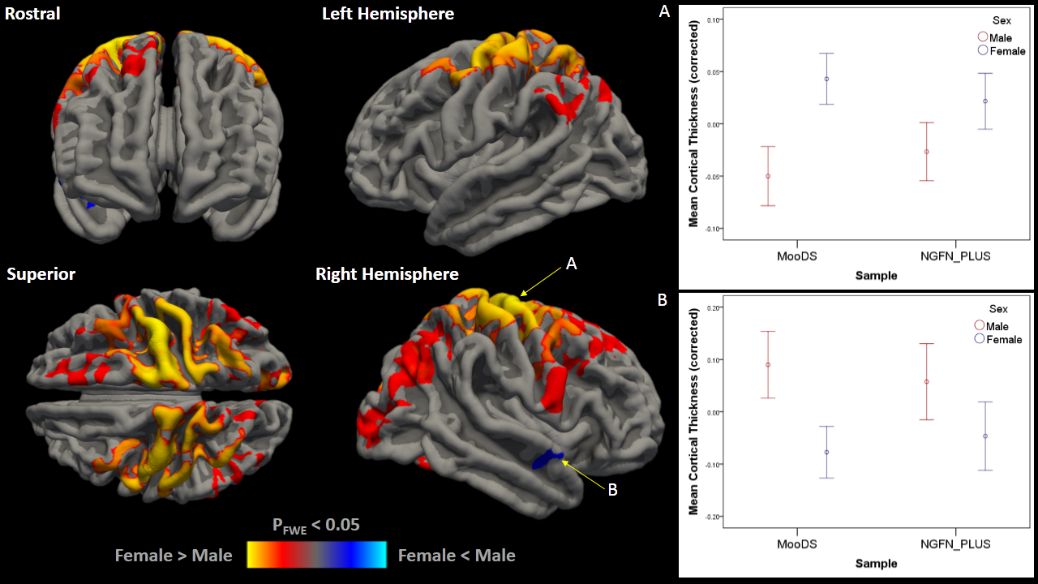
**Supporting Figures**



**S1.** Receiver operating characteristic (ROC) curves at full-width half maximum values ranging from 0 to 3mm (A-D). Within each figure the curves represent adjacency values ranging from the triangular mesh to a geodesic distance of 3mm at the midthickness projected surface. The true positive rate ranges from 0 to 1.0, and the false positive rate ranges from 0 to 0.05. The curves compare the binarized TFCE transformed T-statistic images (PFWE<0.05 threshold) to the classifier score from N=1000 the T-statistic images with random Gaussian noise introduced (mean=0, unit variance), then TFCE transformed and binarized (PFWE<0.05 threshold). The dotted line represents when the true positive rate matches the false positive rate. The area under of the curve (AUC) was greater than 0.97 for all ROC curves.

G:\Tris_Personal_2\FHWM_03B_ROC_various_noise_levels3.tif

**S2.** Receiver operating characteristic (ROC) curves after adding an increasing among of noise at a full-width half maximum of 3mm, and at a geodesic distance of 3mm for the adjacency set used in the TFCE calculation. The curves compare the binarized TFCE transformed T-statistic images (PFWE<0.05 threshold) to the classifier score from N=1000 the T-statistic images with random Gaussian noise introduced (mean=0, unit variance), then TFCE transformed and binarized (PFWE<0.05 threshold). Each curve represent increasing Gaussian noise added from one to three times the standard Gaussian distribution. The dotted line represents when the true positive rate matches the false positive rate. The area under of the curve (AUC) was greater than 0.97 for all ROC curves.



**Figure S3**. The association among cortical thickness and sex across all samples (N=382). There were significant associations among females > males (PFWE<0.05, red-yellow), and females < males (blue-lightblue). Left: the red and blue bars correspond to PFWE < 0.05 and lower. Covariates included age, site, and sample. Right: Mean cluster cortical thickness and 95% confidence intervals by sample. (a) The mean of the right hemisphere cluster in which females > males (6238.7 mm2, max PFWE=0.0008) predicted cortical thickness (partial R2=0.076, F1,374=30.7, p =3.8 x 10-8). (b) The females < male cluster (144.6 mm2, max PFWE =0.041) predicted cortical thickness (partial R2=0.052, F1,374=20.5, p =8.0 x 10-6). Covariates included age, site, and sample.

C:\Users\virtualmachine\Desktop\wincomp_Manuscript_surftfce\vertex_based\FigureS2.tif

**Figure S4**. Sex predicts cortical thickness in (a) MooDS sample (N=199) and (b) NGFN\_PLUS sample (N=183). There were significant associations among females > males (PFWE<0.05, red-yellow), and females < males (blue-lightblue). The bars correspond to PFWE<0.05 and lower. The images are overlaid on the midthickness projected surface. Covariates included age, sex, and site.

G:\Tris_Personal\active_manuscripts\surf_tfce\Manuscript_surftfce\SFigure_IC1_Nback.tif

**Figure S5**. The first independent component (IC1) from the melodic independent analysis of all subjects (N=382). The spatial activation of IC1 fits the typical pattern observed in functional MRI spatial working memory tasks. Area of increased activation (Z scores ranging from 2.3 to 4.0; red-yellow) are observed in the dorsolateral prefrontal cortex and inferior parietal region, and deactivation (Z scores ranging from -2.3 to -4.0; blue-lightblue) are observed in the default mode regions. Activation patterns are projected onto the MNI152\_T1\_2mm\_brain template.

C:\Users\virtualmachine\Desktop\wincomp_Manuscript_surftfce\vertex_based\FigureS4.tif

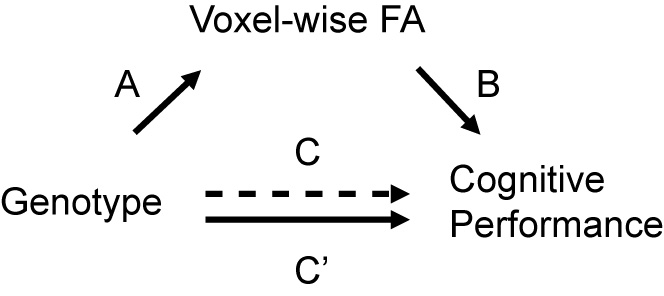
**Figure S6**. The association among cortical surface area and N-Back IC1 in (a) MooDS sample (N=199) and (b) NGFN\_PLUS (N=183). Surface area predicted N-back IC1 across the cortex in both samples (PFWE<0.05, 10000 permutations). The red to yellow bar corresponds to PFWE<0.05 and lower. The images are overlaid on the midthickness projected surface. Covariates included age, sex, and site.

G:\Tris_Personal\active_manuscripts\surf_tfce\Manuscript_surftfce\SFigure_FA_pred_rSupSA_DTI.tif

**Figure S7.** White matter FA predicts the mean right superior frontal surface area from the Desikan-Killiany atlas (PFWE<0.05). The red to yellow bar corresponds to 1-PFWE ranging from 0.05 to 1x10-4, and the green bar corresponds skeleton FA ranging from 0.1 to 0.8. The skeleton is overlaid on the mean FA image from all subjects. Covariates included age, sex, and site.

C:\Users\virtualmachine\Desktop\wincomp_Manuscript_surftfce\vertex_based\FigureS6.tif

**Figure S8**. The mean cluster FA predicted cortical surface area (N=199). The red to yellow bar corresponds to PFWE<0.05 and lower. The images are overlaid on the midthickness projected surface. Covariates included age, sex, and site.



**Figure S9**. Hypothetical voxel-wise mediation model. A = Path A; B = Path B; C = Path C (Total effect), Path C’ (Direct effect), and (C – C’) = indirect effect.

**Supporting Tables**

**Table S1.** Sensitivity, specificity, accuracy, and false discovery are from a simulated effect of known cluster size after various ‘H’ and ‘E’ TFCE settings.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Right Hemisphere (6912 vertices) | | | | Left Hemisphere (4181 vertices) | | | |
|  | Sensitivity | Specificity | Accuracy | FDR | Sensitivity | Specificity | Accuracy | FDR |
| No TFCE | 0.818 | 0.882 | 0.879 | 0.750 | 0.694 | 0.874 | 0.869 | 0.863 |
| H=2, E=1 | 0.798 | 0.984 | 0.975 | 0.293 | 0.666 | 0.955 | 0.947 | 0.702 |
| H=2, E=2/3 | 0.638 | 0.998 | 0.981 | 0.063 | 0.192 | 0.999 | 0.977 | 0.110 |
| H=2, E=0.5 | 0.455 | 0.999 | 0.974 | 0.036 | 0.069 | 1.000 | 0.974 | 0.027 |
| H=1, E=2/3 | 0.740 | 0.989 | 0.978 | 0.231 | 0.484 | 0.977 | 0.963 | 0.626 |
| H=1, E=1 | 0.842 | 0.979 | 0.973 | 0.339 | 0.848 | 0.832 | 0.832 | 0.874 |
| H=4, E=1 | 0.660 | 0.998 | 0.982 | 0.068 | 0.129 | 1.000 | 0.976 | 0.034 |

FDR, false discover rate.

**References**

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C. (2008) Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical image analysis, 12:26-41.

Baron, R.M., Kenny, D.A. (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. Journal of personality and social psychology, 51:1173-82.

Beckmann, C.F., Smith, S.M. (2004) Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE transactions on medical imaging, 23:137-52.

Beckmann, C.F., Smith, S.M. (2005) Tensorial extensions of independent component analysis for multisubject FMRI analysis. NeuroImage, 25:294-311.

Calhoun, V.D., Liu, J., Adali, T. (2009) A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. NeuroImage, 45:S163-72.

Charlet, K., Beck, A., Jorde, A., Wimmer, L., Vollstadt-Klein, S., Gallinat, J., Walter, H., Kiefer, F., Heinz, A. (2014) Increased neural activity during high working memory load predicts low relapse risk in alcohol dependence. Addiction biology, 19:402-14.

Dice, L.R. (1945) Measures of the amount of ecologic association between species. Ecology, 26:297-302.

Esslinger, C., Walter, H., Kirsch, P., Erk, S., Schnell, K., Arnold, C., Haddad, L., Mier, D., Opitz von Boberfeld, C., Raab, K., Witt, S.H., Rietschel, M., Cichon, S., Meyer-Lindenberg, A. (2009) Neural mechanisms of a genome-wide supported psychosis variant. Science, 324:605.

Fischl, B., Dale, A.M. (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences of the United States of America, 97:11050-5.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M. (2002) Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron, 33:341-55.

Fischl, B., Salat, D.H., van der Kouwe, A.J., Makris, N., Segonne, F., Quinn, B.T., Dale, A.M. (2004) Sequence-independent segmentation of magnetic resonance images. NeuroImage, 23 Suppl 1:S69-84.

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

Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M. (1999b) High-resolution intersubject averaging and a coordinate system for the cortical surface. Human brain mapping, 8:272-84.

Greve, D.N., Fischl, B. (2009) Accurate and robust brain image alignment using boundary-based registration. NeuroImage, 48:63-72.

Guo, Y., Pagnoni, G. (2008) A unified framework for group independent component analysis for multi-subject fMRI data. NeuroImage, 42:1078-93.

Jenkinson, M., Bannister, P., Brady, M., Smith, S. (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage, 17:825-41.

Mackinnon, D.P., Warsi, G., Dwyer, J.H. (1995) A Simulation Study of Mediated Effect Measures. Multivariate behavioral research, 30:41.

Oguz, I., Farzinfar, M., Matsui, J., Budin, F., Liu, Z., Gerig, G., Johnson, H.J., Styner, M. (2014) DTIPrep: quality control of diffusion-weighted images. Frontiers in neuroinformatics, 8:4.

Owen, A.M., McMillan, K.M., Laird, A.R., Bullmore, E. (2005) N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Human brain mapping, 25:46-59.

Pruim, R.H., Mennes, M., Buitelaar, J.K., Beckmann, C.F. (2015a) Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. NeuroImage, 112:278-87.

Pruim, R.H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F. (2015b) ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. NeuroImage, 112:267-77.

Smith, S.M. (2002) Fast robust automated brain extraction. Human brain mapping, 17:143-55.

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E. (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. NeuroImage, 31:1487-505.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M. (2004) Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage, 23 Suppl 1:S208-19.

Sobel, M.E. (1982) Asymptotic confidence intervals for indirect effects in structural equation models. Sociological methodology, 13:290-312.

Sobel, M.E. (1986) Some new results on indirect effects and their standard errors in covariance structure models. Sociological methodology:159-186.

Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C. (2010) N4ITK: improved N3 bias correction. IEEE transactions on medical imaging, 29:1310-20.

Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., Nichols, T.E. (2014) Permutation inference for the general linear model. NeuroImage, 92:381-97.