**SUPPLEMENTARY INFORMATION**

**Methods**

***Recruitment***

Thirty-two ADHD boys were recruited in total. Seven boys dropped out of the study due to their dislike of the MRI scanner, 3 were excluded due to co-morbidities (despite the fact we explicitly aimed to recruit only non-comorbid cases), 1 boy did not reach the diagnostic criteria for the combined subtype of ADHD, 1 boy was excluded due to poor task performance and 3 were excluded due to high levels of motion (> 3mm).

Forty–four ASD boys were recruited in total. Of these, 7 boys dropped out of the study due to their lack of tolerance of the MRI scanner, 14 were excluded due to co-morbidities, 1 was excluded due to neurological abnormalities, 2 were excluded due to SSRI use, 1 was excluded due to poor task performance and 2 were excluded due to high levels of motion. Of the remaining ASD boys 13 were diagnosed with Asperger’s and 4 with high functioning Autism.

Thirty–two controls were recruited in total. Of these, 10 were excluded due to high scores on the SDQ and CPRS.

***fMRI Data Analysis Methods***

*Individual Analysis* - fMRI data were first processed to minimise motion related artefacts ([Bullmore *et al.*, 1999a](#_ENREF_2)). A 3D volume consisting of the average intensity at each voxel over the whole experiment was calculated and used as a template. The 3D image volume at each time point was then realigned to this template by computing the combination of rotations (around the x y and z axes) and translations (in x y and z) that maximised the correlation between the image intensities of the volume in question and the template (rigid body registration). Following realignment, data were then smoothed using a Gaussian filter (FWHM, 7.2mm) to improve the signal to noise characteristics of the images. After motion correction, global detrending and spin-excitation history correction, time series analysis for each subject was based on a wavelet-based data resampling method for functional MRI data ([Bullmore *et al.*, 2001](#_ENREF_3), [Bullmore *et al.*, 1999b](#_ENREF_4)). At the individual subject level, a standard general linear modelling approach was used to obtain estimates of the response size (beta) to each N-Back task condition (1-Back; 2-Back; 3-Back) against an implicit baseline (0-Back). Briefly, we first convolved the main experimental conditions (1-Back; 2-Back; 3-Back; all of them contrasted with 0-Back) with two Poisson model functions (peaking at 4s and 8s) after motion correction, global detrending and spin-excitation history correction. We then calculated the weighted sum of these two convolutions that gave the best fit (least-squares) to the time series at each voxel. A goodness-of-fit statistic (the SSQ-ratio) was then computed at each voxel consisting of the ratio of the sum of squares of deviations from the mean intensity value due to the model (fitted time series) divided by the sum of squares due to the residuals (original time series minus model time series). The appropriate null distribution for assessing significance of any given SSQ-ratio was established using a wavelet-based data re-sampling method ([Bullmore *et al.*, 2001](#_ENREF_3)) and applying the model-fitting process to the re-sampled data. This process was repeated 20 times at each voxel and the data combined over all voxels, resulting in 20 null parametric maps of SSQ-ratio for each subject, which were combined to give the overall null distribution of SSQ-ratio. The same permutation strategy was applied at each voxel to preserve spatial correlation structure in the data. Activated voxels, at a <1 level of type I error, were identified through the appropriate critical value of the SSQ-ratio from the null distribution ([Brammer *et al.*, 1997](#_ENREF_1), [Bullmore *et al.*, 1999b](#_ENREF_4)). Individual SSQ-ratio maps were then transformed into standard space, first by rigid body transformation of the fMRI data into a high-resolution inversion recovery image of the same subject, and then by affine transformation onto a Talairach template ([Talairach and Tournoux, 1988](#_ENREF_5)).

*Group Analysis* - A group activation map was produced for each experimental condition (1-Back; 2-Back; 3-Back; contrasted with 0-Back) by calculating the median observed SSQ-ratio over all subjects at each voxel in standard space and testing them against the null distribution of median SSQ-ratios computed from the identically transformed wavelet re-sampled data ([Brammer *et al.*, 1997](#_ENREF_1), [Bullmore *et al.*, 2001](#_ENREF_3)). The voxel-level threshold was first set to 0.05 to give maximum sensitivity and to avoid type II errors. Next, a cluster-level threshold was computed for the resulting 3D voxel clusters. The necessary combination of voxel and cluster level thresholds was not assumed from theory but rather was determined by direct permutation for each data set, giving excellent Type II error control ([Bullmore *et al.*, 1999b](#_ENREF_4)). Cluster mass rather than a cluster extent threshold was used, to minimise discrimination against possible small, strongly responding foci of activation ([Bullmore *et al.*, 1999b](#_ENREF_4)). In all group activation analyses, less than one false positive activation locus was expected for p<0.05 at voxel level and p<0.01 at cluster level.

**Results**

***Group Differences in Clinical Questionnaire Measures***

Multivariate ANOVA showed a significant group effect for all SDQ measures (F(df =10,100)=21 p < 0.0001).Post-hoc analyses showed thatcontrols scored significantly better on all subscales compared to patients (p < 0.01). As expected, ADHD boys scored significantly higher than ASD boys on the conduct and hyperactive/inattentive subscales of the SDQ (p < 0.0001)whileASD boys scored significantly worse than ADHD boys on the peer relations subscale (p < 0.005) and significantlyhigher on the SCQ (F (df =2,50) = 112, p < 0.0001) than ADHD (p < 0.0001) and controls participants (p <0.0001), while ADHD participants scored significantly higher than controls (p <0.0001). As expected, ADHD boys scored higher on the CPRS (F (df =2,52) = 136, p< 0.0001) than ASD (p <0.0001) and controls (p <0.0001) while ASD participants scored higher than controls (p <0.0001).

***fMRI Data - Within-Group Activation Results***

***3Back – 0Back***

*Controls* – During 3Back – 0Back controls activated a bilateral working memory network consisting of bilateral inferior/middle/superior frontal cortices, ACC/SMA, basal ganglia, thalamus, parietal lobe, precuneus, cerebellum/midbrain and left middle temporal lobe.

*ADHD* – While on placebo, the ADHD group activated bilateral inferior/middle frontal cortices, left precentral gyrus, ACC/SMA, bilateral basal ganglia and thalamus, right parietal lobe, precuneus, left middle/superior temporal lobe , left occipital and midbrain/left cerebellum. While on Fluoxetine, they activated bilateral superior/inferior/middle frontal cortex, ACC/SMA, bilateral basal ganglia and thalamus, bilateral parietal lobe, precuneus, right inferior/middle temporal lobe, bilateral midbrain/cerebellar vermis.

*ASD* – While on placebo, the ASD group activated bilateral inferior/middle frontal cortices, left superior frontal cortex, left precentral gyrus ACC/SMA, bilateral basal ganglia and thalamus, bilateral parietal lobe, precuneus, bilateral middle/superior temporal lobe, midbrain/cerebellar vermis and right cerebellum. While on Fluoxetine, the ASD group activated bilateral superior/inferior/middle frontal cortices, left precentral gyrus**,** ACC/SMA, bilateral basal ganglia and thalamus,bilateral insula,bilateral parietal lobe, precuneus**,** left middle/superior temporal lobe, right inferior/middle temporal lobe**,** left middle occipital,midbrain/cerebellar vermis and left cerebellum/fusiform gyrus.

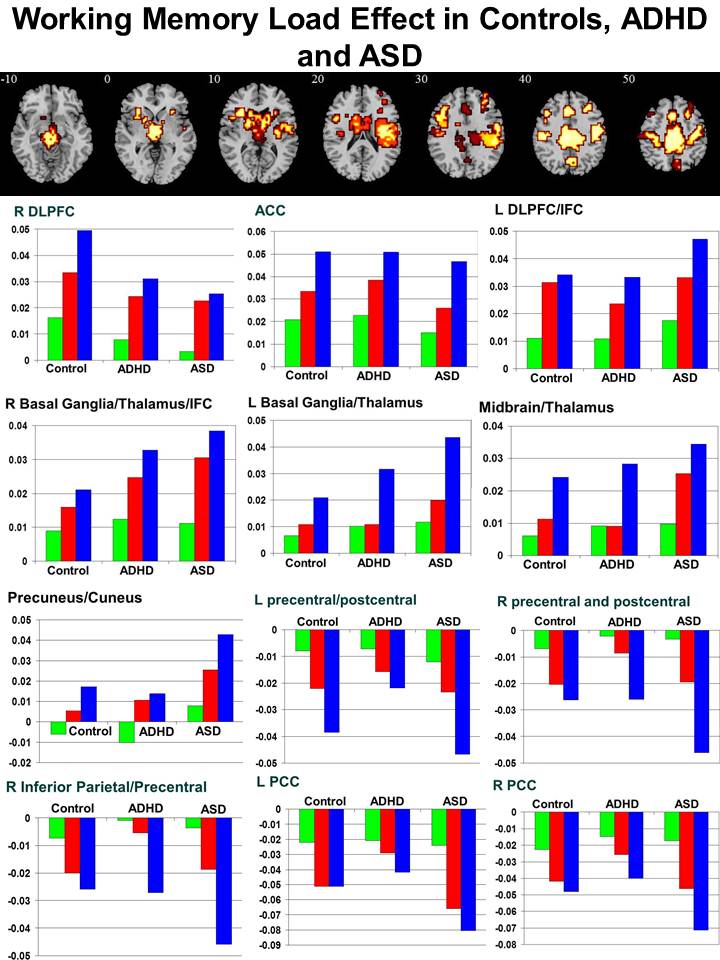
***0-Back – 3-Back***

*Controls*. During 0-Back – 3-Back controls activated a default mode network consisting, bilaterally, of medial prefrontal cortex, precentral and postcentral gyri, putamen and caudate, insula, middle/superior temporal lobe, posterior cingulate cortex, cuneus, fusiform gyrus and right middle occipital lobe.

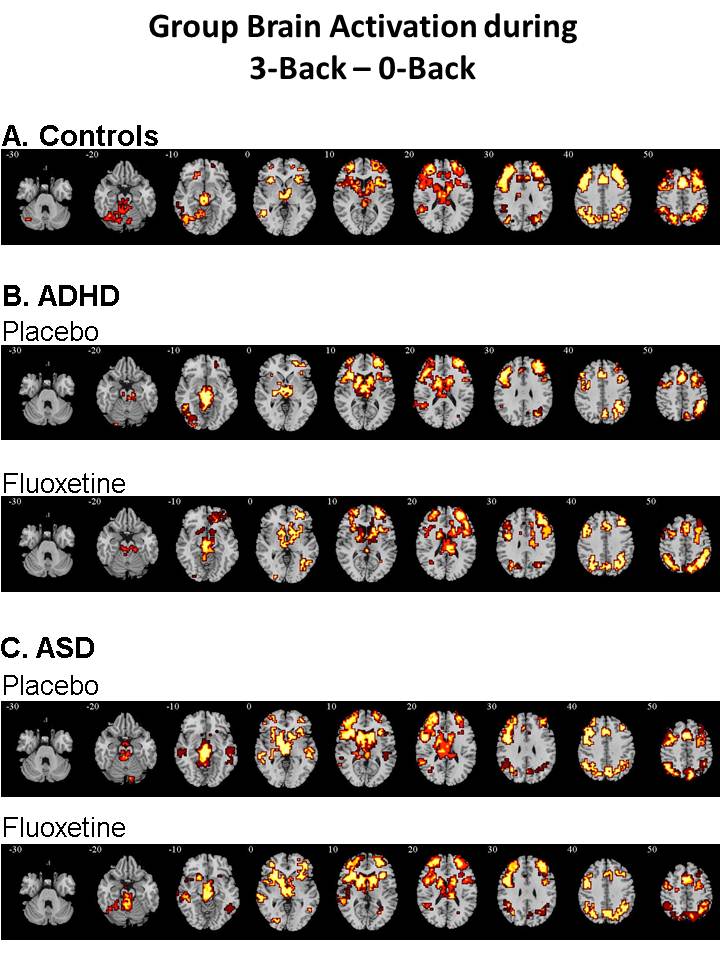
*ADHD*. While on placebo, the ADHD group activated medial prefrontal cortex, left inferior/superior frontal cortex, bilateral precentral and postcentral gyri, left inferior parietal, right middle/superior temporal lobe, bilateral insula/putamen, posterior cingulate cortex, precuneus/cuneus**,** right lingual/fusiform gyrus and right cerebellum. While on Fluoxetine, they activated medial prefrontal cortex, left superior frontal cortex**,** bilateral inferior frontal cortex, bilateral precentral and postcentral gyri**,** bilateral inferior parietal, right middle/superior temporal lobe,bilateral insula/putamen and thalamus,posterior cingulate cortex and precuneus/cuneus.

*ASD* . While on placebo, the ASD group activated medial prefrontal cortex, right superior frontal cortex**,** bilateral precentral and postcentral gyri**,** bilateral inferior parietal lobe,bilateral middle/superior temporal lobe, left putamen, right lentiform nucleus/insula,posterior cingulate cortex,precuneus/cuneus, right fusiform gyrus and cerebellum.While on Fluoxetine, the ASD group activated medial prefrontal cortex,left inferior/superior frontal cortex,bilateral precentral and postcentral gyrus, , left putamen, right lentiform nucleus/insula, bilateral superior temporal lobe, left middle temporal lobe**,** posterior cingulate cortex, precuneus/cuneus, bilateral occipital lobe, bilateral fusiform gyrus and cerebellum.

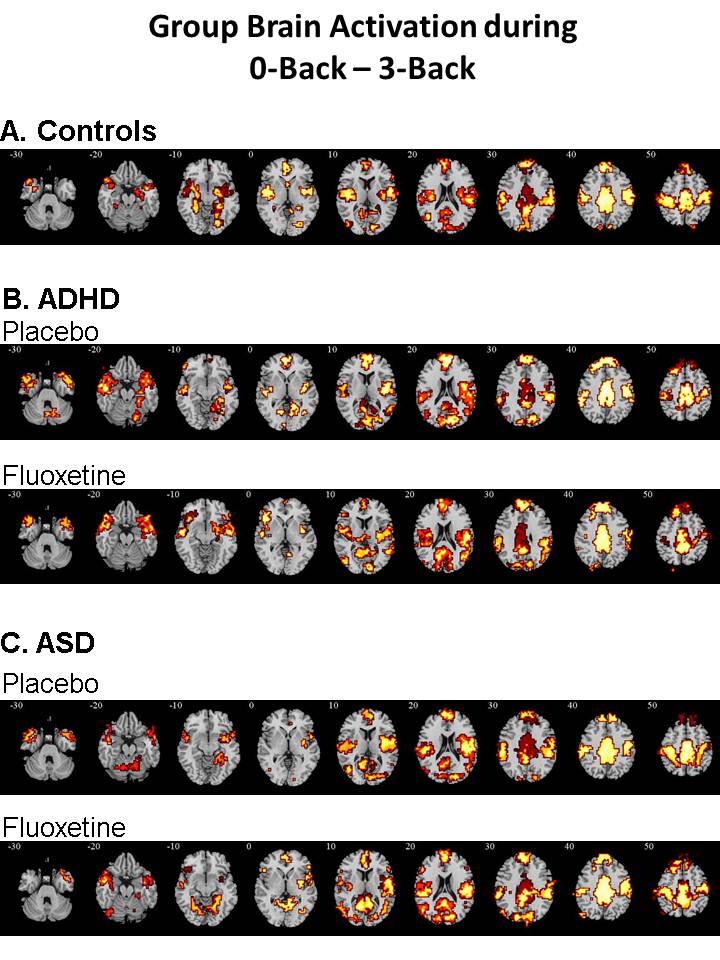
**SUPPLEMENTARY FIGURES**

****

**Supplementary Figure S1** – **Brain activation for working memory load effect in all three groups, i.e. controls, adolescents with ADHD under placebo and adolescents with ASD under placebo. Axial sections for brain activation changes with increasing working memory load across all 3 groups.** Shown underneath are the statistical measures of BOLD response each of the brain regions that showed a significant working memory load effect in all three groups. Green = 1-Back; Red = 2-Back, Blue = 3-Back. Talairach z co-ordinates are indicated for slice distance (in mm) from the intercommissural line. The right side corresponds to the right side of the image.



**Supplementary Figure S2. Within-group activation for the contrast of 3-Back – 0-Back A. Healthy controls, B. Adolescents with ADHD under either placebo or Fluoxetine and C. Adolescents with ASD under either placebo or Fluoxetine**. Axial sections showing within-group brain activation for healthy control boys, boys with ADHD under either placebo or Fluoxetine and boys with ASD under either placebo or Fluoxetine for the contrast of 3-Back – 0-Back. Talairach z co-ordinates are indicated for slice distance (in mm) from the intercommissural line. The right side corresponds to the right side of the image.



**Supplementary Figure S3**. **Within-group activation for the contrast of 0-Back – 3-Back. A. Healthy controls, B. Adolescents with ADHD under either placebo or Fluoxetine and C. Adolescents with ASD under either placebo or Fluoxetine.** Axial sections showing within-group brain activation for healthy control boys, boys with ADHD under either placebo or Fluoxetine and boys with ASD under either placebo or Fluoxetine for the contrast of 0-Back – 3-Back. Talairach z-co-ordinates are indicated for slice distance (in mm) from the intercommissural line. The right side corresponds to the right side of the image.

**References**

**Brammer, M. J., Bullmore, E. T., Simmons, A., Williams, S. C. R., Grasby, P. M., Howard, R. J., Woodruff, P. W. R. & RabeHesketh, S.** (1997). Generic brain activation mapping in functional magnetic resonance imaging: A nonparametric approach. *Magnetic Resonance Imaging* **15**, 763-770.

**Bullmore, E., Brammer, M., Rabe-Hesketh, S., Curtis, V., Morris, R., Williams, S., Sharma, T. & McGuire, P.** (1999a). Methods for diagnosis and treatment of stimulus-correlated motion in generic brain activation studies using fMRI. *Human Brain Mapping* **7**, 38-48.

**Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., Carpenter, T. A. & Brammer, M.** (2001). Colored noise and computational inference in neurophysiological (fMRI) time series analysis: resampling methods in time and wavelet domains. *Human Brain Mapping* **12**, 61-78.

**Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E. & Brammer, M. J.** (1999b). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging* **18**, 32-42.

**Talairach, J. & Tournoux, P.** (1988). *Co-planar stereotaxic atlas of the human brain*. Thieme New York:.