**SUPPLEMENTARY MATERIAL**

**METHODS**

***Analysis of Performance Data***

The free impulsiveness indicator, *k,* can be written as *V=A/(*1*+kD)*, where *V* is the subjective value of reward amount *A*, *D* is the delay and *k* is a constant defining the subject’s rate of discounting with larger k reflecting steeper TD ([Richards *et al.*, 1999](#_ENREF_8)).

***fMRI Image Acquisition***

Gradient-echo echo-planar MR imaging (EPI) data were acquired using the body coil for radio frequency transmission and an 8-channel headcoil for reception. In each of 22 non-contiguous planes parallel to the anterior-posterior commissure, 480 T2\*-weighted MR images depicting BOLD (blood-oxygen level-dependent) contrasts covering the whole brain were acquired with echo time (TE)=30ms, repetition time (TR)=1.5s, flip angle=60 degrees, in-plane voxel size=3.75mm, slice thickness=5mm, slice skip=0.5mm. This EPI dataset provided almost complete brain coverage. Whole-brain high-resolution structural images (inversion recovery gradient echo planar image) used for standard space normalization of individual activation maps were also acquired in the inter-commissural plane with TE=30ms, TR=3s, flip angle=90 degrees, number of slices=43, slice thickness=3.0mm, slice skip=0.3mm, in-plane voxel-size=1.875mm, providing comprehensive brain coverage. Data quality was assured using an automated quality control procedure (Simmons *et al.*, 1999). Most subjects completed 3 additional fMRI tasks in the same session, published elsewhere ([Chantiluke *et al.*, 2014a](#_ENREF_5), [Chantiluke *et al.*, 2014b](#_ENREF_6), [Chantiluke *et al.*, 2014c](#_ENREF_7)). Total scanning time was 1.5 hours.

***fMRI Data Analysis Methods***

*Individual Analysis*

Data were first processed to minimize motion-related artefacts ([Bullmore *et al.*, 1999a](#_ENREF_2)). A 3-D volume consisting of the average intensity at each voxel over the entire 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 and the volume in question and the template (rigid-body registration). Following realignment, data were then smoothed using a Gaussian filter (full-width at half-maximum (FWHM) 7.2 mm in-plane fMRI voxel size) to improve the signal-to-noise ratio of the images ([Bullmore *et al.*, 1999a](#_ENREF_2)). Following motion correction, global detrending and spin-excitation history correction, time series analysis for each subject was conducted based on a previously published wavelet-based resampling method for fMRI 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 of the task conditions (delayed and immediate reward choices) against an implicit baseline. We first convolved the main experimental conditions with 2 Poisson model functions (peaking at 4 and 8s). We then calculated the weighted sum of these 2 convolutions that gave the best fit (least-squares) to the time series at each voxel. A goodness-of-fit statistic (SSQ ratio) was then computed at each voxel consisting of the ration of the sum of squares of deviations from the mean intensity value due to the model (fitted time series) divided by that of the 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 resampled 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 rations for each subject, which were combined to give the overall null distribution of SSQ ratio. This same permutation strategy was applied at each voxel to preserve spatial correlation structure in the data. 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_9)).

*Group Analysis*

A group-level activation map was produced for each group for the experimental condition (delayed-immediate choices) by calculating the median observed SSQ ratios at each voxel in standard space across all subjects and testing them against the null distribution of median SSQ ratios computed from the identically transformed wavelet-resampled 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 dataset, giving excellent type-II error control ([Bullmore *et al.*, 1999b](#_ENREF_4)). Cluster mass rather than a cluster extent threshold was used to minimize discrimination against possible small, strongly responding foci of activation ([Bullmore *et al.*, 1999b](#_ENREF_4)).

**RESULTS**

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

*Controls*

For delayed – immediate choices, controls activated dorsomedial prefrontal cortex (dmPFC) andanterior cingulate cortex (ACC), bilateral anterior insula, pre and postcentral gyri, inferior parietal lobe (IPL), occipital lobe and bilateral cerebellum. They showed more activation to immediate relative to delayed choices in dmPFC extending into ventromedial prefrontal cortex (vmPFC), left inferior frontal cortex (IFC), bilateral caudate and thalamus, left medial temporal lobe and hippocampus, left IPL and bilateral precuneus extending into the posterior cingulate cortex (PCC) (Fig S1A).

*ADHD patients on placebo*

Under placebo, ADHD patients activated the ACC, bilateral pre and postcentral gyrus, bilateral occipital lobe and PCC and bilateral cerebellum more to delayed choices while they activated more for immediate choices left IPL, right insula, left middle frontal gyrus and right pre and postcentral gyrus and IPL (Fig S1B).

*ADHD patients on fluoxetine*

Under fluoxetine, ADHD patients activated the right dorsolateral prefrontal cortex (DLPFC), ACC, right IFC and insula extending into caudate, putamen and ventral striatum, left middle temporal lobe, bilateral IPL, occipital lobe and cuneus and bilateral cerebellum more to delayed choices and activated bilateral inferior temporal lobes, bilateral cerebellum and occipital lobe, superior parietal lobe, dmPFC and ACC and right postcentral gyrus, posterior and IPL more to immediate choices (Fig S1C).

**REFERENCES**

**Brammer, M. J., Bullmore, E. T., Simmons, A., Williams, S. C. R., Grasby, P. M., Howard, R. J., Woodruff, P. W. R. & Rabe-Hesketh, 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. *Medical Imaging, IEEE Transactions on Medical Imaging* **18**, 32-42.

**Chantiluke, K., Barrett, N., Giampietro, V., Brammer, M., Simmons, A., Murphy, D. G. & Rubia, K.** (2014a). Inverse Effect of Fluoxetine on Medial Prefrontal Cortex Activation During Reward Reversal in ADHD and Autism. *Cerebral Cortex*.

**Chantiluke, K., Barrett, N., Giampietro, V., Brammer, M., Simmons, A. & Rubia, K.** (2014b). Disorder-dissociated effects of fluoxetine on brain function of working memory in attention deficit hyperactivity disorder and autism spectrum disorder. *Psychological Medicine* **45**, 1195-1205.

**Chantiluke, K., Barrett, N., Giampietro, V., Santosh, P., Brammer, M., Simmons, A., Murphy, D. & Rubia, K.** (2014c). Inverse fluoxetine effects on inhibitory brain activation in non-comorbid boys with ADHD and with ASD. *Psychopharmacology*, 1-12.

**Richards, J. B., Zhang, L., Mitchell, S. H. & Wit, H.** (1999). Delay or probability discounting in a model of impulsive behavior: effect of alcohol. *Journal of the Experimental Analysis of Behavior* **71**, 121-143.

**Talairach, J. & Tournoux, P.** (1988). *Coplanar stereotaxic atlas of the human brain, a 3-dimensional proportional system: an approach to cerebral imaging*. Thieme: New York.

**Figure S1**. Group brain activation maps

**(A) Controls**



**(B) ADHD Placebo**



**(C) ADHD Fluoxetine**



**Supplementary Figure S1.** Group activation maps. Axial slices showing within-group brain activation for the contrasts of delayed-immediate reward choices (red) and immediate-delayed reward choices (blue). (A) healthy controls, (B) ADHD patients under placebo and (C) ADHD patients under fluoxetine for Talairach z-coordinates are indicated for slice distance (in mm) from the intercommissural line. The right side of the brain corresponds to the right side of the image.