



Supporting Online Material for

Orbitofrontal Dysfunction in Patients with Obsessive-Compulsive Disorder and Their Unaffected Relatives

Samuel R. Chamberlain,* Lara A. Menzies, Adam Hampshire, John Suckling, Naomi A. Fineberg, Natalia del Campo, Mike Aitken, Kevin Craig, Adrian M. Owen, Edward T. Bullmore, Trevor W. Robbins, Barbara J. Sahakian

*To whom correspondence should be addressed. E-mail: srchamb@gmail.co

Published 18 July 2008, *Science* **321**, 421 (2008)
DOI: 10.1126/science.1154433

This PDF file includes:

Materials and Methods
Figs. S1 to S7
Table S1

Supporting Online Material

Supplemental Methods

Recruitment and screening

OCD patients were recruited via the National OCD Treatment Service (England and Wales), Queen Elizabeth II Hospital, Welwyn Garden City, Hertfordshire, United Kingdom. Patients gave written permission for a first-degree relative to be contacted (preferably a similarly aged same sex sibling). Healthy controls were recruited via media advertisements. All potential recruits undertook a clinical interview including the Mini International Neuropsychiatric Inventory (MINI) to screen for current axis-I disorders¹. Mood status was assessed using the Montgomery-Asberg Depression Rating Scale (MADRS)², and verbal IQ was quantified using the National Adult Reading Test (NART)³. Patients were recruited on the basis of mainly washing/checking symptoms without hoarding, and freedom from axis-I comorbidities (including current depression, i.e. freedom from fulfillment of DSM-IV criteria and MADRS total score < 16). First-degree relatives and comparison subjects were excluded if they had DSM-IV axis-I disorders (including major depressive disorder, OCD itself, and other anxiety disorders). OCD symptom severity was measured using the Yale-Brown Obsessive Compulsive Scale (Y-BOCS)^{4, 5}. All participants were right-hand dominant

according to the Edinburgh Handedness Inventory⁶, and reported no history of head trauma, neurological disease, or contraindications for MRI. Participants provided written informed consent prior to participation and the study was approved by the local research ethics committee.

Task description

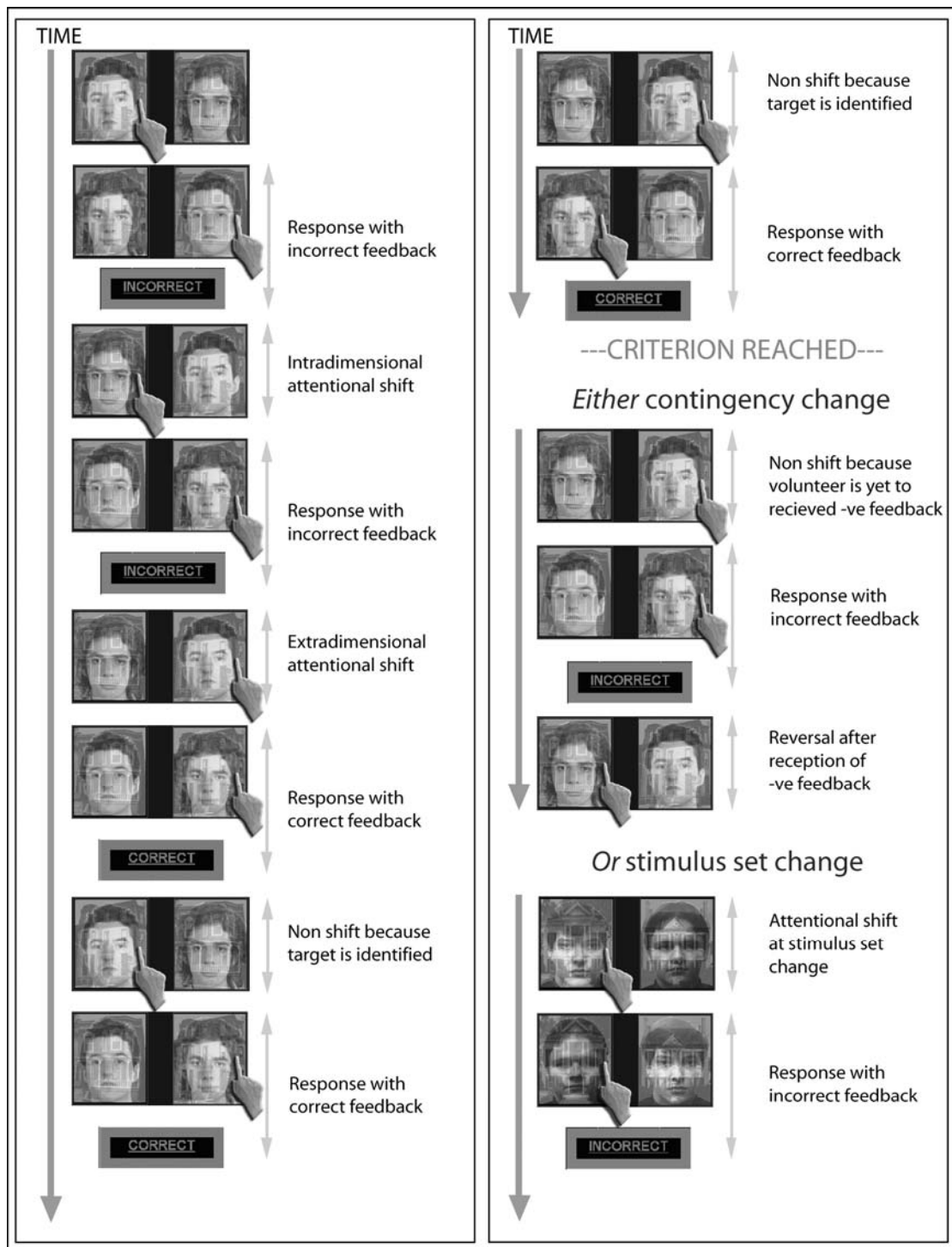
An attentional switching task was used⁷ in which participants attempted to work out which object was the target in a stimulus set consisting of two faces and two buildings. The stimulus set was presented as two compound object pairs appearing on the left and right of the screen. Both compound object pairs consisted of a face and a building superimposed on top of each other. Each stimulus approximately subtended a visual vertical angle of 6 degrees and a horizontal angle of 6.2 degrees, with a total combined horizontal angle of 15 degrees. On each trial, participants indicated which side of the screen they believed the target was on, using a button box. After every second response, feedback was presented on the screen for 0.6 seconds, indicating whether the object they had chosen was the target or not (“CORRECT” or “INCORRECT”).

Fig. S1 illustrates a typical series of trials. In this example, the participant initially chooses the face in the left superimposed face/building pair and so indicates left with the button box by pressing button with right index finger. When the response is made, the stimuli are removed from the screen and reappear after a short interval rearranged with the chosen face on the right of

the screen superimposed on the other building; the participant therefore indicates right by pressing button with right middle finger. Because the face-building combinations swap from one trial to the next, the program can compute which item was selected and because (in this example) it is not the target, negative feedback ('INCORRECT') is given. Subsequently the stimuli reappear on the screen and the participant selects the other face (*intra-dimensional, ID shift*). Following the second response, negative feedback is given and the participant switches to select the building on the right of the screen (*extra-dimensional, ED shift*). Following the second response to the building, positive feedback ('CORRECT') is given because the participant has correctly identified the target item. When the stimuli reappear on the screen, the participant responds to the same building, as they now know that it is the target (early correct response). They receive positive feedback on the second response, and so continue to select the same building (late correct response). After responding correctly again they receive positive feedback and have now reached the criterion of six correct responses in a row. One of two things then happens; either a new stimulus set is presented, in which case the participant starts searching for the new target (*set change*). Alternatively, the reward contingency changes, in which case the participant responds twice more to the same building (because they have no way of knowing that anything has changed) before receiving negative feedback. They must then inhibit their responses to the recently rewarded target object, and start trying to identify which of the other three possible items has become the target (*reversal*). It is important to note that the extra-dimensional and intra-dimensional shift events, along with the feedback, do not always occur in the sequence shown

because the order in which the stimuli are tested is determined entirely by the choices made by participants.

Fig. S1: the flexibility task



Imaging Acquisition

Scanning was undertaken at the Wolfson Brain Imaging Centre using a 3 Tesla Siemens scanner with 21 slices (4mm slices with 1mm inter-slice gap) per image and a temporal resolution (TR) of 1.1 seconds and in plane resolution of 3.2 x 3.2 mm. 850 T2-weighted echo-planar images, depicting BOLD contrast were acquired per experimental block, and the first 18 were discarded to avoid T1 equilibrium effects. Images were slice time acquisition corrected, reoriented, subject motion corrected, spatially normalised to the standard Montreal Neurological Institute echoplanar imaging template, smoothed with an 8mm full-width at half-maximum Gaussian kernel, and modelled using Statistical Parametric Mapping 5 (SPM 5, <http://www.fil.ion.ucl.ac.uk/spm>). We used the default smoothing option in SPM (8mm) as per our previous validation study⁷. For optimal sensitivity, filter size should be approximately the same size as the signal to be detected (matched filter theorem); since the cortex is in the order of a few mm thick, a smoothing kernel in the order of mm rather than cm is appropriate.

The time series were high pass filtered, and the haemodynamic response was modelled to the stimulus onsets and durations. For switch events durations were up until the time of response at which stage the stimuli were removed from the screen, whereas for feedback events durations were from stimulus

onset up until the point of removal of the subsequent feedback from the screen.

The experimental in-scanner acquisition itself consisted of two 15-minute sessions. As the timing was response driven, the number of switches completed varied for each participant. The inter-stimulus-interval was randomly jittered from 0.6 to 1.6 seconds. Prior to undertaking the scanning, participants were pre-trained on the behavioral task, to minimize the likelihood of between-group behavioral differences occurring that would have confounded interpretation of the fMRI data. In a previous behavioral (out-of-scanner) pilot study, patients with OCD ($n = 20$) exhibited reversal deficits versus healthy controls ($n = 20$) [patients mean \pm SD 8.38 ± 4.00 errors; controls 4.97 ± 1.47 ; $F(1,38)=12.81$, $p < 0.005$]. Responses were made using the index and middle fingers of the right hand on a button box to indicate left and right, respectively, and were recorded throughout the experimental acquisition.

Event modeling

The event modeling focused on individual types of participant response and these were defined according to which objects were currently and previously selected. The four event types of interest in this paper were: reversal, set-change, intra-dimensional switch, and extra-dimensional switch. Responses when the target was known on the basis of prior positive feedback, were also

included, and were divided into the first (early) and subsequent (late) correct responses (at the early correct responses, an important behavioral change occurred as the participant stopped trying to work out which object was the target). Finally, the responses with positive and negative feedback events were included separately for the periods of time when the target identity was being derived, and when the target identity was known. Contrasts were generated corresponding to brain activation for *working out* solutions (solution search minus knowing the target), *reversal* (reversals at contingency change minus shift at set change), and *ED-shifting* (ED minus ID attention shifts)⁷.

Statistical approach

In-scanner error data on the task were analysed using repeated-measures ANOVA (within subject factors: intra-dimensional versus extra-dimensional; non-reversal versus reversal; between subject factor: group). Response times for different events of interest were analysed using one-way ANOVAs (between subject factor: group). Significant effects of interactions were explored via least significant difference (LSD) tests. Significance was set at $p < 0.05$.

Neural regions activated during each of the three contrasts of interest were first identified by conducting a second-level group analysis in SPM 5 across all 39 subjects (main effect of condition, False Discovery Rate (FDR) corrected, $p < 0.05$). Between-group differences in brain activation within these regions were then investigated using non-parametric permutation analysis

implemented in Cambridge Brain Activation software (CAMBA v2.0.0) (www-bmu.psychiatry.cam.ac.uk/software/)^{8, 9}. Specifically, activations for each contrast of interest for each subject were each entered into voxelwise 3x1 ANOVAs in CAMBA, with group (patients, relatives, controls) as the between-subject factor. Non-parametric permutation analysis was used to identify significant clusters in which there was a main effect of group. To provide stringent control for multiple comparisons, statistical correction was applied such that the expected number of false positive identified clusters per whole brain map of interest was less than one. Significant clusters were described in terms of their maximal MNI co-ordinates and the Brodmann Areas (BA) they subtended. Mean activation data from these identified clusters were then extracted and subjected to non-parametric Monte Carlo testing in SPSS v15, in order to compare activation between individual groups (patients versus controls; relatives versus controls; patients versus relatives) ($p < 0.05$ uncorrected).

Supplemental results

Demographic and clinical characteristics

As can be seen in table S1 below, the three groups did not differ significantly in terms of age, NART IQ, sex ratio, or handedness. The mean MADRS scores for all groups were well beneath clinically significant depressive mood, although there was a non-significant trend towards groups differing on total MADRS scores. There were significant group differences in terms of Y-BOCS scores: patients exhibited scores consistent with moderate severity OCD while all controls and relatives demonstrated zero scores on this instrument. Nine patients were receiving stable doses of serotonin reuptake inhibitors (SRIs) at the time of participation (doses/day: sertraline: 50mg, 100mg, 200mg, 200mg; citalopram: 20mg, 20mg, 40mg; fluoxetine: 80mg; fluvoxamine 300mg). Mean activation from brain clusters showing significant effects of group, for each contrast of interest, did not correlate significantly with symptom severity in the patients (all $p > 0.30$, Pearson's r).

Table S1: Demographic and clinical characteristics of the sample

Variable	Patients (n=14)		Relatives (n=12)		Controls (n=15)		F (df=2,40)	p
	Mean	SD	Mean	SD	Mean	SD		
Age (years)	31.7	10.8	39.5	11.4	34.8	8.7	1.870	0.17
NART IQ	114.6	5.3	113.5	8.0	115.8	6.9	0.394	0.68
Males:Females	2:12		2:10		2:13			>0.5 ^a
Handedness	91.43	24.1	89.1	14.5	95.3	15.5	0.372	0.69
MADRS	5.79	7.8	2.8	3.9	1.0	3.6	2.787	0.07
Y-BOCS-o	11.92	3.3	0.0	0.0	0.0	0.0	176.270	< 0.01 *
Y-BOCS-c	12.46	3.1	0.0	0.0	0.0	0.0	218.773	< 0.01 *
Y-BOCS-t	23.62	6.8	0.0	0.0	0.0	0.0	162.515	< 0.01 *
* p<0.05, significant difference between groups								
^a chi-squared test								

Behavioral results

In terms of errors made on the task, there was a significant effect of reversal [$F(1,37) = 13.311, p < 0.01$] and a significant effect of set-shift on errors [$F(1,37) = 53.538, p < 0.01$]. However there was no significant effect of group and no interactions obtained significance (all $p > 0.10$). Thus, as can be seen in Fig. S2, participants in general made more errors when reversing responses than not; and when shifting attention between rather than within dimensions.

For response times (Fig. S3), one-way ANOVAs revealed significant group differences for the ED shift, ID shift, reversal, and set-change stages only [least significant $F(2,37) = 3.334, p < 0.05$]. This was due to significantly lengthened response times in relatives versus patients and controls, as can be seen in the graph (all $p < 0.05$); patients did not differ significantly from control on these measures (all $p > 0.10$).

Fig. S2: In-scanner errors on the task

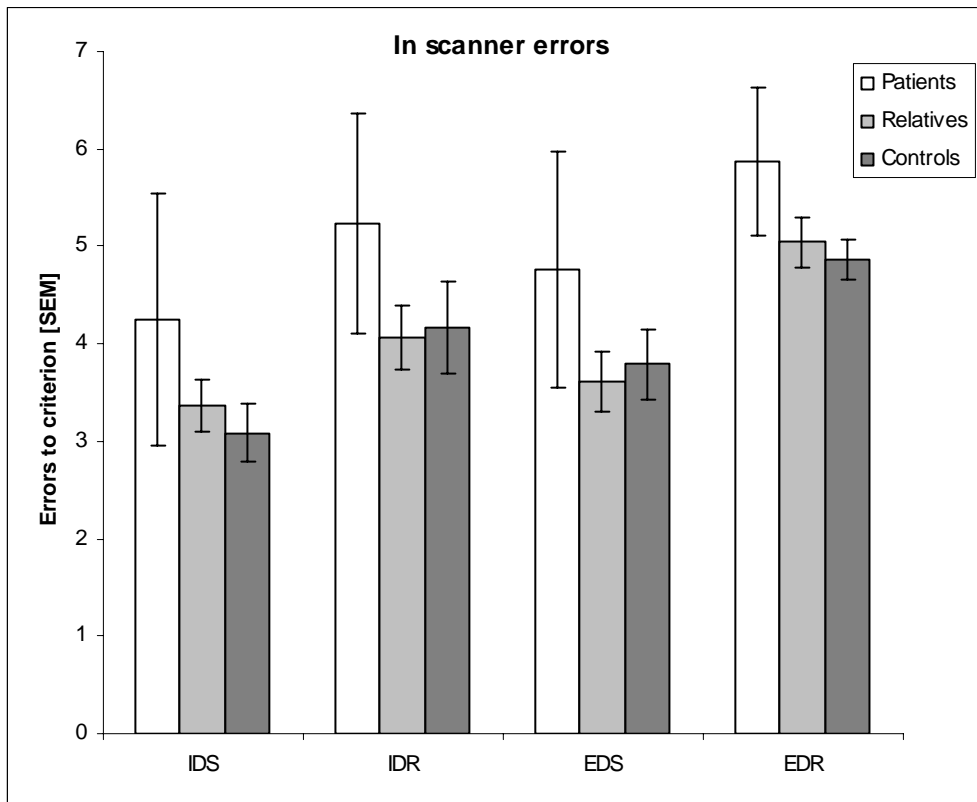
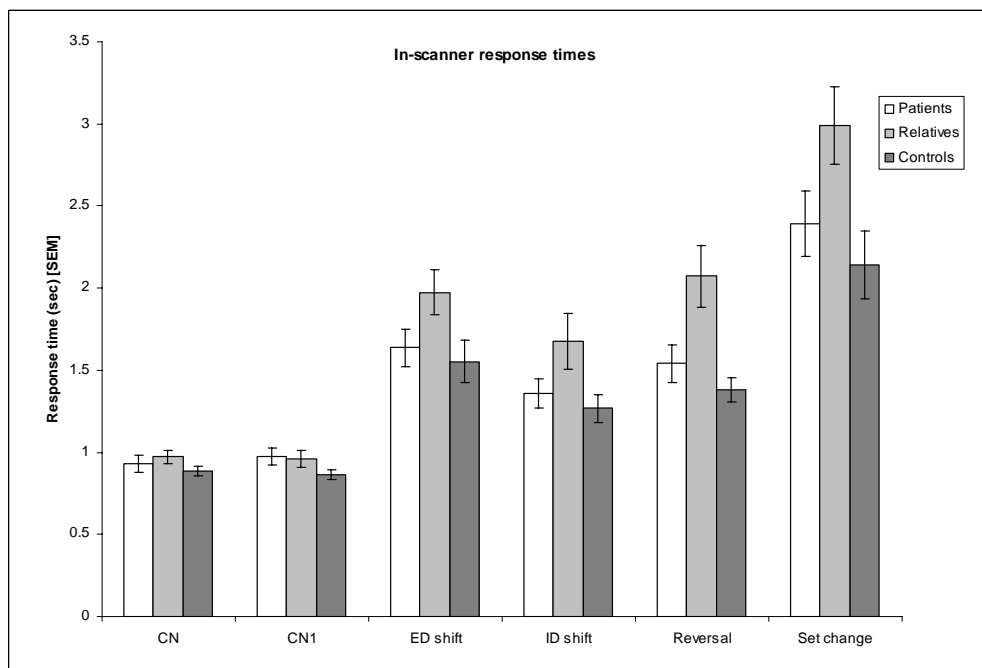


Fig. S3. In-scanner response times on the task



fMRI results

Neural regions significantly activated when *working out* solutions and during *reversal* are indicated in Fig. S4. CAMBA analysis showed that there was a main effect of group (patients, relatives, controls) when *working out* solutions in four clusters as follows: cluster 1 (4623 voxels): left sub-thalamic nucleus, bilateral calcarine, bilateral lingual, left occipital, and bilateral fusiform cortices; and left pallidal, bilateral thalamic, and right cerebellar regions [maximum: $x=-6$, $y=-14$, $z=-4$] (BA 17, 18, 19, 30, 37); cluster 2 (474 voxels): left OFC and lateral PFC [maximum $x=-46$, $y=44$, $z=-4$] (BA 10, 45, 46, 47); cluster 3 (466 voxels): right OFC and lateral PFC [maximum $x=52$, $y=34$, $z=16$] (BA 10, 45, 46, 47); cluster 4 (341 voxels): left parietal lobe [maximum $x=-44$, $y=-28$, $z=50$] (BA 2, 3, 40). Monte Carlo tests showed significant under-activation in all four clusters in patients versus controls (all $p < 0.05$), and in relatives versus controls (all $p < 0.05$). Patients did not differ significantly from relatives in any cluster (all $p > 0.10$).

CAMBA analysis showed that there was a main effect of group in four clusters during *reversal*: cluster 1 (182 voxels): left OFC and left lateral PFC [maximum $x=-42$, $y=50$, $z=-2$] (BA 10, 11, 46, 47); cluster 2 (84 voxels): right OFC and right lateral PFC [maximum $x=40$, $y=56$, $z=0$] (BA 10, 11, 46, 47); cluster 3 (123 voxels): left parietal lobe [maximum $x=-32$, $y=-48$, $z=50$] (BA 40); cluster 4 (93 voxels): right parietal lobe [maximum $x=50$, $y=-40$, $z=46$] (BA 40). Monte Carlo tests showed significant under-activation in all clusters in patients versus controls and in relatives versus controls (all $p < 0.05$, see Fig.

1). Patients did not differ significantly from relatives in any cluster (all $p > 0.10$).

Brain activation during ED-shifting did not withstand FDR correction and thus was not investigated at the between-group level in CAMBA.

Since the primary analysis was restricted to those regions significantly activated by the task across all subjects overall, we also conducted a supplementary between-group analysis with CAMBA across the whole brain without masking at the cluster-level (statistical threshold of less than one false positive cluster per map). This showed that the right OFC finding during reversal was robust even at this very stringent threshold; it also confirmed no other previously unidentified regions in which the groups differed (Fig. S5). Graphs of power based on the region of OFC identified in this analysis allow sample sizes to be estimated for future studies at false positive rates (alpha) of 0.05, for patients compared to controls (Fig. S6) and relatives of patients compared to controls (Fig. S7).

Fig. S4. Brain activation when working out solutions (A) and during reversal learning (B), across all subjects. Tables indicate peak activation co-ordinates in MNI space, and brains depict all regions of significant activation (FDR corrected $p < 0.05$).

A.

voxel p(FDR-cor)	voxel T	voxel equivZ	MNI x,y,z {mm}	Region	BA
<0.001	10.52	7.16	-26 -62 50	Left parietal	7
<0.001	8.56	6.35	16 -68 52	Right parietal	7
<0.001	8.53	6.34	42 -54 -14	Right posterior	
<0.001	8.6	6.37	-46 8 32	Left frontal	9
<0.001	8.2	6.18	-4 20 46	Left cingulate	32
<0.001	7.8	5.99	-28 6 58	Left frontal	6



B.

voxel p(FDR-cor)	voxel T	voxel equivZ	MNI x,y,z {mm}	Region	BA
0.005	5.95	4.97	28 48 -4	Right cingulate	10
0.005	5.77	4.86	24 40 -4	N/A	
0.011	4.57	4.05	26 62 2	Right frontal	10
0.005	5.8	4.88	-30 56 0	Left frontal	10
0.005	5.66	4.79	-38 52 -2	Left frontal	
0.005	5.57	4.73	8 -60 50	Right parietal	7
0.014	4.34	3.89	-36 -48 50	Left parietal	40
0.017	4.23	3.8	-40 -44 44	Left parietal	40
0.007	5.02	4.37	42 -54 42	Right parietal	39
0.019	4.14	3.74	46 -44 48	Right parietal	40
0.020	4.06	3.68	40 32 36	Right frontal	9
0.036	3.58	3.3	40 20 44	Right frontal	8
0.021	4	3.63	34 10 54	Right frontal	6
0.029	3.74	3.43	8 18 4	Right caudate	

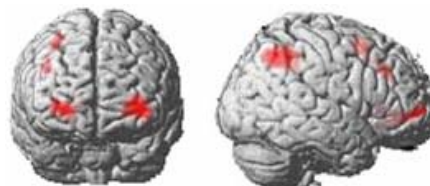


Fig. S5. Permutation analysis across the whole brain (without mask restriction) confirmed a significant effect of group on brain activation in a single cluster (997 voxels) comprising the right OFC and right lateral PFC (blue) [maximum x=44, y=44, z=16] (BAs 10,11, 46, 47). Numbers refer to z co-ordinates in MNI space. Graph, patients and relatives showed significant under-activation in this cluster versus controls (* $p < 0.05$, *** $p < 0.001$, Monte Carlo tests).

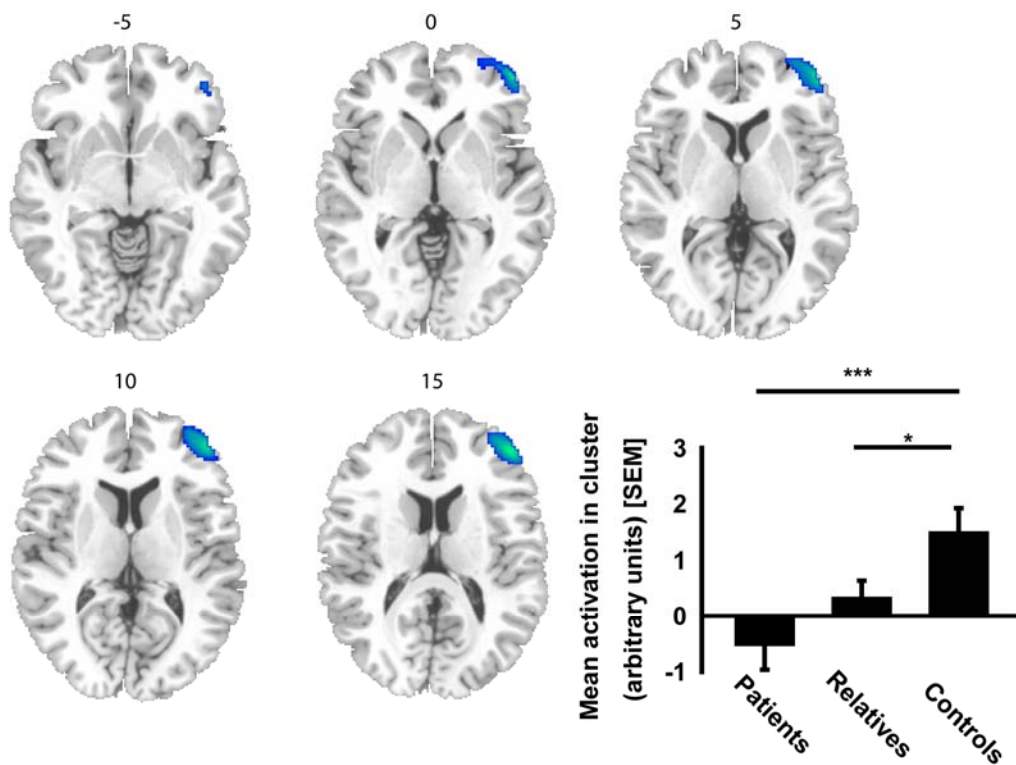


Fig. S6. Statistical power (1-beta) to detect a significant difference in brain activation during reversal between patients and controls, based on data from the whole brain analysis described in Fig. S5 and parametric assumptions, alpha = 0.05 (graph created using GPower software¹⁰).

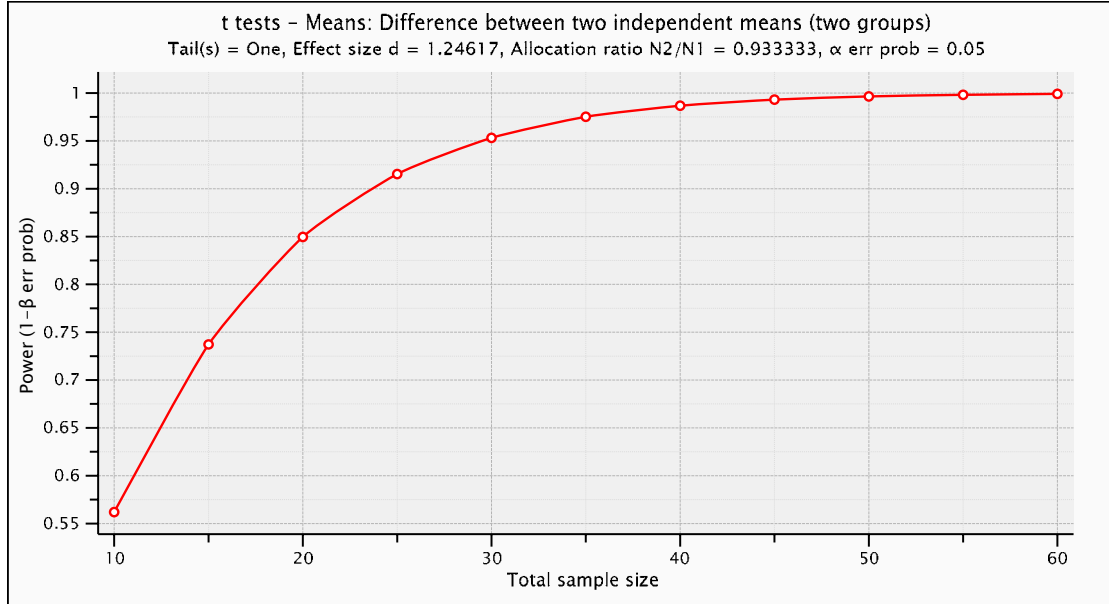
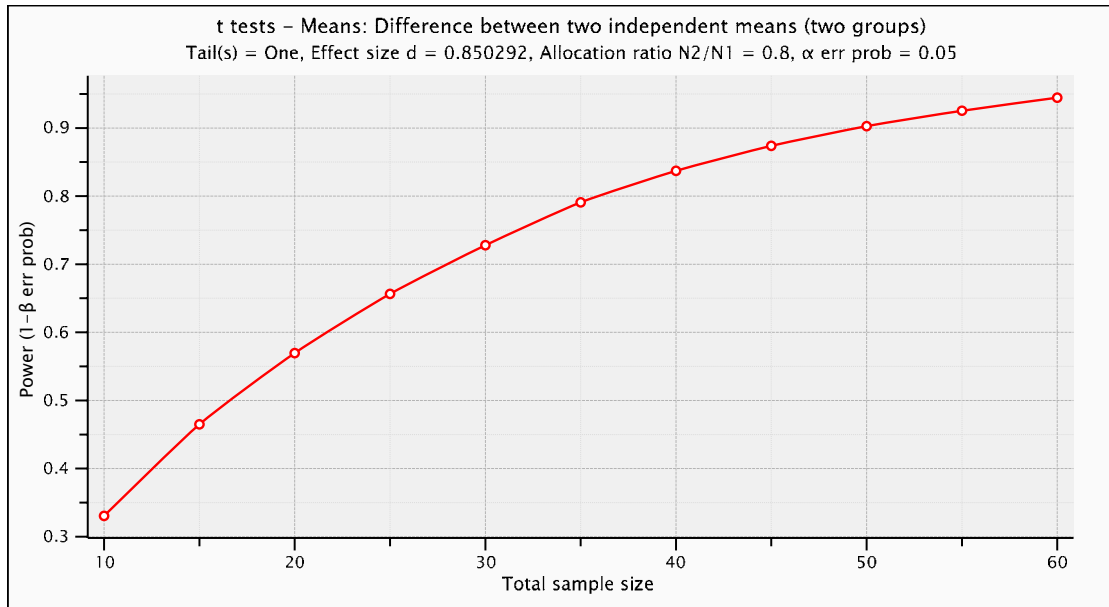


Fig. S7. Graph showing statistical power (1-beta) to detect a significant difference in brain activation during reversal between patients' relatives and controls, based on data from the whole brain analysis described in Fig. S5 and parametric assumptions, alpha = 0.05 (graph created using GPower software¹⁰).



Supplemental Discussion

Most individuals with OCD exhibit substantial symptom overlap across multiple domains¹¹. In order to maximize homogeneity within the study, we limited the sample to those patients with prominent (but not exclusively) archetypal washing/checking symptoms without hoarding. Our study was not designed to compare imaging results across different OCD symptom dimensions.

This study was neither designed nor powered to delineate the effects of SRI medication on brain activation, and we deemed it ethically inappropriate to withdraw treatment from those patients who were taking SRIs, for the purposes of this study. Of course this does mean that any quantitative comparison between OCD patients and relatives is confounded by medication – however, this is not a major thrust of the paper, which is to show the qualitative similarity between these two groups. Patients and relatives showed similar brain abnormalities (i.e. parietal and OFC hypoactivation in the reversal contrast, Fig. 1). Thus, the novel strategy of including relatives allowed us to demonstrate that brain dysfunction exists in the absence of medication confounds, in those at risk of OCD.

This experiment was sufficiently powered to detect group effects on task-related changes in signal – our key a priori hypothesis was confirmed at a stringent statistical threshold. Power calculations are not easily applicable to the non-parametric approach. Precise power analyses with available tools

depend upon parametric statistical assumptions that are not necessarily met. Furthermore, ours is the first study to measure brain activation in unaffected relatives of patients with OCD. We have conducted power analyses based on the most conservative finding in this study, that of right OFC under-activation in patients and relatives in the whole brain analysis without masking (Fig. S5). During reversal, patients showed mean activation of $-0.525 \pm \text{SD } 1.638$ arbitrary units, and controls 1.500 ± 1.628 ($n=15$, $n=13$). This corresponded to a large effect size (Cohen's d) of 1.242. Relatives showed mean activation of 0.330 ± 1.066 arbitrary units ($n=12$), Cohen's $d = 0.850$. Thus this study was amply powered to confirm our key finding with large effect sizes. Power calculations (Fig. S6, Fig. S7) are provided for future studies. From these figures it can be seen that this study had ~90% power to detect under-activation in patients versus relatives, and ~70% power to detect significant under-activation in relatives versus controls, in the right OFC. These likely reflect under-estimates since the power of non-parametric permutation analysis is likely to be greater than that of parametric approaches.

The recruitment of non-depressed patients represents a key strength of this study. Relatives of patients showed similar hypo-activation abnormalities to patients, yet did not differ significantly from controls on MADRS scores (post-hoc t-test, $p = 0.39$). Thus sub-clinical depression did not contribute to the findings. We also re-ran the imaging analysis with MADRS scores as a covariate and the findings were essentially unchanged. The issue of how clinical depression affects performance on a similar task is largely an

independent one; in general there is not a simple effect of depression upon performance on the ID/ED shift test.

There is quite frequently a mismatch between behavioral and imaging deficits that ultimately has to be explained in mechanistic terms. In this case of OCD, it has previously been shown by us that reversal learning is dependent on the integrity of the orbitofrontal cortex in both animal and human neuropsychological and imaging studies. It is likely that latent deficits in performance can often be compensated by other mechanisms (e.g. verbal mediation), especially for a relatively simple test such as reversal learning. In fact, as it happens the version of the reversal test we used in this study DID elicit a deficit in performance in OCD patients when not pre-trained probably because of features that have yet to be analysed between this test and previous ones (e.g. the nature of the discriminative stimuli). The important fact is that even when performance on the test has been equated by practice there is a residual deficit apparent in the OFC as revealed by the BOLD response in the OFC. Moreover, this result is consistent with a wealth of other evidence of OFC abnormality in OCD¹².

With respect to extra-dimensional shifting, a deficit has also been reported in OCD by using an ID/ED test¹³ which again is somewhat different from the one employed here, which did not show impairment on initial examination and therefore which may be less behaviorally sensitive. Furthermore, activation in the ED-shifting contrast across all subjects did not withstand correction for multiple comparisons in this study. The possibility of functional impairments in

both the OFC and the IFC is supported by the results of a parallel structural imaging study of OCD which also addressed the issue of endophenotype, and identified grey matter deficits in both regions in patients and relatives, although with a different neuropsychological test and in the structural rather than functional imaging domain (the Stop-signal reaction time test¹⁴).

With respect to the failure to observe deficits in other tests associated with OFC dysfunction, such as decision-making and gambling tasks – it is true that OCD patients exhibit no major impairments in these domains. However, this could be for similar reasons as for the unimpaired reversal learning. The hypothesis could readily be tested with a functional imaging version of the gambling task. Another issue is precisely which sector of the OFC mediates performance on the reversal and gambling tasks – it is possible that these differ for medial versus lateral OFC.

References

1. D. V. Sheehan *et al.*, *J Clin Psychiatry* **59 Suppl 20**, 22 (1998).
2. S. A. Montgomery, M. Asberg, *Br J Psychiatry* **134**, 382 (May, 1979).
3. H. E. Nelson. (Windsor, NFER-Nelson., 1982).
4. W. K. Goodman *et al.*, *Arch Gen Psychiatry* **46**, 1012 (Nov, 1989).
5. W. K. Goodman *et al.*, *Arch Gen Psychiatry* **46**, 1006 (Nov, 1989).
6. R. C. Oldfield, *Neuropsychologia* **9**, 97 (Mar, 1971).
7. A. Hampshire, A. M. Owen, *Cereb Cortex* (Jan 25, 2006).
8. E. T. Bullmore, J. Suckling, S. Overmeyer, S. Rabe-Hesketh, E. Taylor, M. J. Brammer, *IEEE Trans Med Imaging* **18**, 32 (Jan, 1999).
9. J. Suckling, E. Bullmore, *Hum Brain Mapp* **22**, 193 (Jul, 2004).
10. E. Erdfelder, F. Faul, A. Buchner, *Behav Res Methods, Instruments, & Computers* **28**, 1 (1996).
11. D. Mataix-Cols, M. C. Rosario-Campos, J. F. Leckman, *Am J Psychiatry* **162**, 228 (Feb, 2005).
12. L. Menzies, S. R. Chamberlain, A. R. Laird, S. M. Thelen, B. J. Sahakian, E. T. Bullmore, *Neurosci Biobehav Rev* (Oct 17, 2007).
13. S. R. Chamberlain, N. A. Fineberg, A. D. Blackwell, T. W. Robbins, B. J. Sahakian, *Am J Psychiatry* **163**, 1282 (Jul, 2006).
14. L. Menzies *et al.*, *Brain* **130**, 3223 (Dec, 2007).