The orbitofrontal cortex is necessary for temporal context memory

Audrey Duarte\textsuperscript{1,2}, Richard N. Henson\textsuperscript{1}, Robert T. Knight\textsuperscript{3}, Tina Emery\textsuperscript{1}, and Kim S. Graham\textsuperscript{1,4}

\textsuperscript{1}Medical Research Council - Cognition and Brain Sciences Unit, Cambridge, UK
\textsuperscript{2}School of Psychology - Georgia Institute of Technology, Atlanta, Georgia, USA
\textsuperscript{3}Department of Psychology and the Helen Wills Neuroscience Institute, University of California, Berkeley, USA
\textsuperscript{4}Wales Institute of Cognitive Neuroscience and School of Psychology, Cardiff University, UK

Abstract

Lesion and neuroimaging studies suggest that the orbitofrontal cortex (OFC) supports temporal aspects of episodic memory. However, it is unclear whether OFC contributes to the encoding and/or retrieval of temporal context and whether it is selective for temporal relative to non-temporal (spatial) context memory. We addressed this issue with two complimentary studies: functional magnetic resonance imaging (fMRI) to measure OFC activity associated with successful temporal and spatial context memory during encoding and retrieval in healthy young participants and a neuropsychological investigation to measure changes in spatial and temporal context memory in OFC lesion patients. Imaging results revealed that OFC contributed to encoding and retrieval of associations between objects and their temporal but not their spatial contexts. Consistent with this, OFC patients exhibited impairments in temporal but not spatial source memory accuracy. These results suggest that the OFC plays a critical role in the formation and subsequent retrieval of temporal context.

INTRODUCTION

Episodic memory has been defined as the recollection of an event defined by a unique spatial and temporal context (Tulving, 1985). These spatial and temporal details, along with other contextual information such as perceptual attributes and internally generated feelings, fall under the rubric of “sources” assessed in source memory tasks. In a typical source memory experiment, different contexts are implemented during the study or “encoding” phase and subsequently during test or “retrieval”, participants may be asked to judge in which of the study contexts the item was previously encountered. Although memory for these different contexts may be highly integrated at a phenomenological level, there is evidence to suggest that somewhat dissociable brain regions may contribute to memory for spatial and temporal contextual details.

Numerous human neuroimaging have implicated medial temporal lobe (MTL) structures in spatial context memory (Spiers & Maguire, 2007 for review) and results from lesion studies in rats (Kesner & Hopkins, 2006 for review), non-human primates (e.g. Alvarado, Wright, &
Bachevalier, 2002; Malkova & Mishkin, 2003)) and humans (Kessels, de Haan, Kappelle, & Postma, 2001 for review) further suggest that the MTL may be necessary for successful spatial memory. Somewhat less is known about the neural basis of temporal context memory. For example, across the few imaging studies that have compared memory for temporal and non-temporal (i.e. spatial) information, (Fujii et al., 2004; Hayes, Ryan, Schnyer, & Nadel, 2004; Henson, Shallice, & Dolan, 1999; Nyberg et al., 1996; Rekkas et al., 2005), there is little consensus as to what brain regions are particularly associated with temporal context memory. Some lesion evidence also supports a role for the MTL in humans (Downes, Mayes, MacDonald, & Hunkin, 2002; Kesner & Hopkins, 2001), though not as notably as for spatial memory. Focal lesions to the prefrontal cortex (PFC) in humans have more often been associated with impairments in memory for temporal context relative to both item recognition (Milner, Corsi, & Leonard, 1991; Shimamura, Janowsky, & Squire, 1990) and spatial context memory (Kopelman, Stanhope, & Kingsley, 1997). The lesions in these studies, however, could either not be determined (Shimamura et al., 1990), or were large and heterogeneous (Kopelman et al., 1997; Milner et al., 1991), making it difficult to determine what specific PFC regions were involved.

A more targeted approach that distinguishes between different PFC regions may prove useful in investigating temporal context memory. For example, patients with focal orbitofrontal cortex (OFC) damage can exhibit difficulty distinguishing between currently-relevant and currently-irrelevant memory traces, so-called “temporal context confusions” (Dalla Barba, 1993; Schneider & Ptak, 1999). Patients typically remember the content of events but they often misattribute events or aspects of events from one time with those of another. This impaired temporal processing has been offered as a mechanism for the confabulation sometimes observed in patients with OFC damage, although temporal context confusions also occur in the absence of confabulation (Gilboa et al., 2006). Experimentally, temporal processing has been assessed in these patients using a continuous recognition task, with repeated presentation of items across and within task runs (Gilboa et al., 2006; Schneider & Ptak, 1999; Schneider, Treyer, & Buck, 2000; Treyer, Buck, & Schneider, 2003). In these studies, however, memory for temporal context was measured indirectly and patients’ deficits may have been related to proactive interference, rather than temporal context memory, per se.

The temporal context confusions observed in OFC lesion patients have been proposed to reflect impaired retrieval processes, such as deficient monitoring and evaluation of retrieval cues (Gilboa et al., 2006). However, human imaging (Frey & Petrides, 2000, 2002; Ranganath, Heller, Cohen, Brozinsky, & Rissman, 2005) and single unit recording studies in non-human primates (Rolls, Browning, Inoue, & Hernadi, 2005) have implicated the OFC in the encoding of new information and activity in this region may predict successful subsequent memory performance (Frey & Petrides, 2003; Ranganath et al., 2005). Thus, it is possible that the OFC may contribute to both encoding and retrieval of temporal contextual information. It is difficult to determine from patient studies alone, however, whether anterograde memory deficits are related to impairments in encoding, retrieval or both. Furthermore, even relatively focal lesions may encompass several subregions, limiting one’s ability to make anatomically-specific conclusions. Thus, it remains unclear which specific regions of the OFC, if any, are necessary for successful encoding and/or retrieval of temporal context.

In order to more directly assess the role of the OFC in the formation and retrieval of temporal context associations, we conducted two complementary experiments using the same paradigm: collecting functional magnetic resonance imaging (fMRI) data from healthy individuals in Experiment 1 and behavioral data from patients with well-characterized OFC lesions in Experiment 2. Together, these experiments provide complementary information concerning the anatomical basis of episodic memory for temporal context. Specifically, during study, pictures of objects were presented in one of two spatial (location) and temporal (study list)
contexts. During test, participants saw studied and unstudied pictures, and judged which they had seen previously. While the fMRI data provided anatomical specificity, and information about the involvement of OFC in successful encoding and retrieval, the patient data provided evidence about the necessity of this region to successful context memory.

**METHODS**

**Experiment 1 - fMRI**

**Participants**—15 young adults (7 women; 23 years of age (range 18–30); 15.1 years of education (range 13–19)) were recruited from local universities, science fairs and the Medical Research Council Cognition and Brain Sciences Unit volunteer panel. Participants were paid for their time and signed consent forms approved by the Cambridge Local Research Ethics Committee. Participants were right-handed, fluent English speakers with normal or corrected-to-normal vision (using MRI-compatible glasses when necessary). None of the participants reported cognitive complaint, a history of psychiatric or neurological disorder or psychoactive drug use. All MRI scans were screened by a radiologist for abnormalities (malformation, hydrocephalus, etc.).

**Procedure**—Stimuli consisted of 384 grayscale line drawings of nameable concrete objects subtending a maximum vertical and horizontal visual angle of up to 4.16°. A short practice version of the experiment was administered to participants outside of the scanner immediately prior to scanning. Both study and test periods were scanned. Participants responded using buttons on a box placed under their right hand.

There were 128 trials at study, split into two lists (“sets”) that were separated by a 5 minute MPRAGE scan. This separation was to make the study lists temporally more distinct. Half of the objects were presented above a central fixation cross and half were presented below. Objects were presented for 1500 ms in one of 16 possible vertical positions along the midline, with 8 above and 8 below fixation, given that piloting showed this was effective in reducing spatial source accuracy performance from ceiling (and producing a close match to temporal source accuracy). In order to encourage incidental encoding of the spatial/temporal context, participants performed a semantic judgment task on each object, responding whether it would, or would not, fit inside a shoebox. Study trials were separated by a 1500 ms fixation screen.

Study was followed by 4 test sessions of 64 studied objects (32 from each study set, half of which previously presented above fixation and half previously presented below) plus 32 unstudied items, presented in a pseudorandom order. Participants first made a subjective judgment of whether they "remembered" the object from the study phase, whether it seemed "familiar" and therefore likely to have come from the study phase, or whether it was a new (unstudied) object. Detailed instructions for the appropriate use of these “remember”, “familiar” and “new” response categories was based on previous studies (Gardiner & Java, 1991; Rajaram, 1993), where we used the term “familiar” in place of “know” to ease exposition. Objects were all centrally presented above a response cue stating these 3 choices for 3 seconds. After a 500 ms fixation screen, a new response cue appeared for 3 seconds, asking the participants to make an objective source decision about either the temporal or spatial context.

Participants performed either the spatial or temporal retrieval task for blocks of 24 trials within each test session. In the spatial blocks, participants decided where the object was presented on the screen during the study phase (“top” or “bottom”); in the temporal blocks, they decided in which study set the object was presented (“set 1” or “set 2”). A third response option of “don’t know” was offered when the relevant context could not be recollected. For all “new” judgments, participants were instructed to respond “don’t know” to the second response cue, in order to balance the number of responses across all conditions. A second 500 ms fixation screen was
presented after the objective source decision and before the next trial. A full analysis of the subjective measures of recollection are presented elsewhere (Duarte, Graham, & Henson, 2008; Duarte, Henson, & Graham, 2008) and for the purposes of the present manuscript, we collapsed objective (spatial and temporal source) decisions across their associated subjective (remember and familiar) decisions. The Huynh-Feldt correction, reflected in the p-values, was used in the behavioral analyses, where appropriate. Two-tailed T-tests were used for pairwise comparisons of the behavioral data.

**fMRI acquisition**—Scanning was performed on a 3T Siemens TIM Trio system. Functional data were acquired using a gradient-echo pulse sequence (32 transverse slices oriented along the anterior-posterior commissural axis, repetition time 2 s, echo time 30 ms, 3 × 3 × 3.5 mm voxels, 0.8 mm interslice gap). Each encoding session (n = 2) included 193 volumes and each retrieval session (n = 4) included 356 volumes. The first 5 volumes per session were discarded to allow for equilibration effects. A map of the magnetic field was also acquired in order to correct for inhomogeneity-related distortions in the functional EPI scans. A high-resolution T1-weighted magnetization-prepared rapid-acquisition gradient echo (MPRAGE) image was collected for anatomical localization.

**fMRI analysis**—Data were analyzed using SPM2. EPI Images were unwarping using the individual’s magnetic field map to undistort the EPI images (Hutton et al., 2002; Jezzard & Balaban, 1995) and subsequently realigned. Distortions in EPI images can be particularly problematic in the OFC, due to its location near the frontal sinuses. The unwarping corrects for potential voxel displacement of observed activations in the OFC and other brain regions. Furthermore, because of susceptibility to signal dropout in the OFC, which cannot be corrected post acquisition, we carefully inspected the local maxima on the individual EPI images to ensure that they were not located within regions of dropout. The resulting mean EPI image from realignment was used to estimate normalization parameters to the standard MNI EPI template, which were then applied to all EPI volumes. Normalized images were resliced to 3 × 3 × 3 mm and smoothed with an 8mm full-width half-maximum isotropic Gaussian kernel. The data were high-pass filtered to a maximum of 1/128 Hz and grand mean scaled to 100.

Statistical analysis was performed in two stages. In the first stage, neural activity was modeled by a sequence of delta functions at onset of the various trial types and convolved with a canonical hemodynamic response function. These trial types were hits (i.e. remember and familiar) for each source, studied items incorrectly judged to be new (i.e. “misses”), correct rejections to unstudied items and incorrect responses to unstudied items (i.e. “false alarms”). The timecourses were downsampled to the middle slice to form the covariates for the General Linear Model. Temporal autocorrelations within a session were corrected using an AR(1) model. For each participant and session, 6 covariates representing residual movement-related artifacts, determined by the spatial realignment step, were included in the first level model to capture residual (linear) movement artifacts, including movement-by-distortion interactions.

Contrasts of the parameter estimates for each participant were submitted to the second stage of analysis (treating participants as a random-effect). Separate ANOVA models were created for study and test periods. The four contrasts corresponded to 1) hits with correct spatial source, 2) hits with “don’t know” spatial source, 3) hits with correct temporal source, and 4) hits with "don’t know" temporal source, conforming to a 2x2 factorial analysis of context retrieval success by type of context (temporal/spatial). Incorrect source judgments were not included in this analysis, given that it is less clear what cognitive processes contributed to these judgments. For the study phase, each contrast was against a baseline of “new” responses subsequently made to studied items (“misses”); for the test phase, it was against a baseline of “new” responses made to unstudied items (“correct rejections”). Using one event-type (e.g. misses) as a baseline is necessary in designs like this in which the absence of explicit interstimulus intervals means
that overall levels of activity across event-types are not estimated efficiently (Josephs & Henson, 1999). A weighted least squares estimation procedure was used to correct for inhomogeneity of covariance across the four contrasts.

Region of interest (ROI) analyses were conducted using bilateral masks for the orbitofrontal gyri (superior, medial, middle, inferior, gyrus rectus) from the Automatic Anatomical Labeling (AAL) of the MNI brain. The ROIs are displayed on the MNI reference brain in Figure 1(A). First, task (spatial, temporal) X accuracy (correct source, don’t know source) interactions were evaluated, correcting for multiple comparisons across voxels within each ROI using Random Field Theory (RFT) to implement Small-Volume Correction (SVC) under a Family-Wise Error (FWE) correction of $p < 0.05$. Simple effects analyses (within task) to explore the source of the interactions were similarly evaluated for these ROIs. To give an idea of the extent of activation, the number of voxels within these ROIs is reported at an uncorrected alpha of $p < 0.001$. All contrasts were one-tailed $t$-tests. Neural activity, plotted for the local maxima within these ROIs, reflected the parameter estimates for the event-related regressors and had arbitrary units.

**Experiment 2 - Patients**

**Participants**—Seven patients with PFC lesions (1 right unilateral; 6 bilateral), and fourteen age- (48.3 ± 18.8 years), education- (15.0 ± 1.8 years) gender- (4 male) and handedness- (14 right) matched controls participated. Two controls were selected to be directly matched on each of these criteria (age, education and handedness) to one patient. Control participants were recruited from the Medical Research Council Cognition and Brain Sciences Unit volunteer panel. None of the controls reported cognitive complaint, a history of psychiatric or neurological disorder (including depression and epilepsy), vascular disease (including diabetes) or psychoactive drug use. All participants were paid for their time and signed consent forms approved by the Cambridge Local Research Ethics Committee and the Institutional Review Board of the University of California, Berkeley.

Patients were identified by review of MRI scans from the outpatient populations at the Veterans Affairs Medical Center (VAMC) in Martinez, CA and the Alta Bates Medical Center in Berkeley, CA and Addenbrookes hospital in Cambridge, UK. Patients were included if they were at least 6 months post-cerebral accident and had no history of any other medical, neurological or psychiatric disorder. All patients were in stable condition at the time of participation. None of the patients exhibited confabulation or language difficulties. MRicro software was used to overlay patient lesions onto the MNI template and a Brodmann area map that was matched to this template was then used to approximate the areas of lesion overlap. A neurologist (RTK) examined each of the patient’s lesion locations and grouped patients according to previously outlined architectonic divisions (Petrides & Pandya, 1994).

Individual patient demographics can be found in Supplemental Table 1. The average lesion volume was 33 cm$^3$ and was centered in the gyrus rectus, medial and superior OFC (Brodmann areas 11 and 10) for all patients, with varying degrees of damage in the middle and inferior orbitofrontal gyr (Brodmann area 47). Lesion overlap for the patient group is shown in Figure 1(B). Lesion volume and Brodmann areas for individual patients are listed in Supplemental Table 1 and individual patients’ lesion reconstructions are shown in Figure 2. Patients were carefully selected so that damage was localized to the OFC and did not impinge upon other frontal areas, including the lateral PFC. Given the variability in the volume of the lesions across patients, we investigated the correlations between volume and source memory accuracy. Neither the lesion volume correlation with spatial nor with temporal source memory accuracy estimates were reliable [$r$'s < 0.2, $p$'s > 0.68].
In order to ensure that the patients did not exhibit profound cognitive deficits, such as amnesia, that would impair their ability to perform the experimental task, patients and their matched controls older were administered a battery of standardized neuropsychological tests. The battery included tests of long-term verbal memory, verbal fluency, working memory span, and executive function: Rey Auditory Verbal Learning Test (RAVLT) (Rey, 1941), the Controlled Oral Word Association Test ("FAS") (Benton, Hamsher, & Sivan, 1983), the Wisconsin Card Sorting Test (Lezak, 1995), the Wechsler Memory Scale-Revised (WMS-R) Digit Span Forward and Backward (Wechsler, 1997) and the Rey Complex Figure Test (Rey, 1941). The results from these tests are presented in Table 1. Individual patient neuropsychological test results are presented in Supplemental Table 2. There were no significant differences between the groups for any of these tests \( t(19)’s < 1 \).

Procedure—The procedure was the same as that of Experiment 1 except for a few changes implemented to ensure that the patients could perform the tasks satisfactorily. This included splitting the experiment into two study-test replications, and making the test phases self-paced (the study phases had the same timing as in Experiment 1). Thus each of the two study sets contained 32 trials, separated by a 5 minute break. After the first study session, participants performed one self-paced spatial and one self-paced temporal block, the order of which was counterbalanced across participants but held constant between each patient and matched-control pair. After a period of neuropsychological testing (see above), the same study-test procedure was repeated (with new stimuli), balancing the order of the test sessions within participant (e.g., Study1: Set1 - 5 minute break - Study1: Set2 - Test1: Spatial - Test1: Temporal - Neuropsychological testing - Study2: Set1 - 5 minute break - Study2: Set2 - Test2: Temporal - Test2: Spatial). For group comparisons, ANOVAs that yielded significant interactions were followed up with planned contrasts (t-tests).

RESULTS

Experiment 1 - fMRI

Behavioral results—The mean proportions of correct, incorrect and don’t know source judgments to studied items, and of new judgments made to studied (i.e. misses) and unstudied items (i.e. correct rejections), are shown in Table 2. Item recognition accuracy was estimated by the Pr measure of discriminability, i.e. \( p(\text{hits}) - p(\text{false alarms}) \). These estimates were 72.9% and 72.5% for spatial and temporal retrieval blocks respectively, with no significant difference between the two \( t(14) < 1 \). Nor was there a difference in the proportion of studied items given a correct source judgment for spatial versus temporal contexts \( t(14) < 1 \). However, there was a greater tendency for incorrect source judgments for temporal than spatial contexts (with a corresponding reduction in "don't know" responses) \( t(14)’s > 2.89, p’< 0.01 \). To accommodate this potential difference in response bias, source accuracy was also estimated by Pr, excluding "don’t knows", i.e., \( Pr = p(\text{correct}) - p(\text{incorrect}) \). These estimates were 34% and 20% for spatial and temporal contexts respectively. Analysis confirmed that source Pr was greater for spatial than temporal contexts \( t(14) = 2.9, p = 0.01 \), though both were significantly greater than chance (0%) \( t(14)’s > 6.5, p’< 0.0001 \).

Mean RTs for the item recognition judgments are shown in Table 2. A 2x2 ANOVA employing factors of Context (spatial, temporal) and Source Response (correct, "don’t know") yielded a main effect of Source Response \( F(1, 14) = 176.2, p < 0.0001 \) and no other effects. Pairwise contrasts confirmed that correct source responses were faster than "don’t know" source responses for both tasks \( t(14)’s > 9.76, p’< 0.0001 \). Similarly, the same ANOVA performed for the source memory judgment RTs, shown in the table, yielded a significant main effect for Source Response \( F(1, 14) = 176.2, p < 0.0001 \). Pairwise contrasts confirmed that correct source responses were faster than "don’t know" source responses for these RTs as well \( t(14)’s > 9.76, p’< 0.0001 \).
Finally, an ANOVA performed for the study phase RTs for items subsequently associated with correct source and don’t know spatial and temporal source judgments, shown in Table 3, revealed a significant main effect of Source Response \([F(1, 14) = 11.5, p = 0.004]\). Pairwise contrasts confirmed that subsequent don’t know source responses were faster than subsequent correct source responses for both tasks \([t(14)]'(s > 1.9, p’s < 0.05]\).

**fMRI results**—To identify regions associated with successful encoding of context, we compared the mean event-related response at test to studied items that were subsequently recognized and associated with correct source judgments with that to studied items that were subsequently recognized but given a "don’t know" response for the source. To identify regions associated with successful item encoding, we averaged the mean event-related responses across these two categories at study. To identify regions associated with successful retrieval of context, we compared the mean event-related response at test to studied items that were recognized and associated with correct source judgments with that to studied items that were recognized but given a "don’t know" response for the source. To identify regions associated with successful item recognition, we averaged the mean event-related responses across these two categories at test.

During study, there was a reliable interaction between subsequent source memory accuracy and context in the right medial orbitofrontal gyrus ROI \([x = 9, y = 54, z = -12, T = 3.20, p < 0.05 SVC-corrected, 8 voxels]\) and the right gyrus rectus ROI \([x = 6, y = 33, z = -24, T = 3.24, p < 0.05 SVC-corrected, 10 voxels]\). As shown for the right medial orbitofrontal region in Figure 3, activity was greater for temporal than spatial source memory in these regions, with greater activity for correct than don’t know temporal source trials and no reliable difference between correct and don’t know spatial source trials \([maximum T = 1.3, minimum SVC-corrected p = 0.9]\). No OFC ROI demonstrated greater activity for subsequent spatial source memory accuracy or source accuracy effects that were reliable for both contexts \([maximum T = 1.4, minimum SVC-corrected p = 0.92]\). Finally, no OFC ROI showed reliable effects of item encoding \([maximum T = 2.3, minimum SVC-corrected p = 0.62]\). 1

During test, no region revealed reliable interactions between source memory accuracy and context at the corrected threshold. There was a marginal interaction in the left gyrus rectus ROI \([x = -18, y = 27, z = -15, T = 2.84, p = 0.1 SVC-corrected]\). As shown in Figure 3, activity was greater for don’t know than correct temporal source trials \([T = 3.36, p = 0.03 SVC-corrected, 6 voxels]\) with no reliable difference between correct and don’t know spatial source trials \([maximum T = 0.59, minimum SVC-corrected p = 0.98]\). No OFC ROI showed greater activity associated with spatial source retrieval accuracy or source accuracy effects that were reliable for both contexts \([maximum T = 1.7, minimum SVC-corrected p = 0.80]\). Successful item retrieval for both spatial and temporal tasks was associated with reliable activity in bilateral inferior orbitofrontal gyrus \([x = -42, y = 42, z = -6, T = 4.33, 19 voxels; x = 30, y = 24, z = -6, T = 3.78, 4 voxels, p’s < 0.05 SVC-corrected]\), with greater activity for correctly recognized (correct + don’t know source) items than correctly rejected new items. No OFC ROI exhibited significant differences between spatial and temporal item retrieval \([maximum T = 2.0, minimum SVC-corrected p = 0.5]\). 2

---

1ROI analyses conducted for a bilateral mask for the hippocampus from the Automatic Anatomical Labeling (AAL) of the MNI brain revealed a reliable interaction between subsequent item memory accuracy (source correct + don’t know) > misses) and context (spatial, temporal) in the left hippocampus \([-24, -24, -12; T = 3.52, p = 0.03 SVC-corrected]\), with greater activity for spatial than temporal item memory at study. The effect sizes (partial eta squared) for this hippocampal interaction and for the right medial orbitofrontal source accuracy X context interaction identified at study were 0.59 and 0.23, respectively. This suggests that although memory accuracy effects were obtained in both orbitofrontal and hippocampal regions at study, at least some of the results in the orbitofrontal cortex may be a bit more subtle than effects within the hippocampus.
Experiment 2 – Patients

The mean proportions of correct, incorrect and don’t know source judgments to studied items, and of new judgments made to studied items (i.e. misses) and unstudied items (i.e. correct rejections), are shown in Table 4. The Pr estimates of item recognition were 84.0% and 84.7% for spatial and temporal retrieval blocks, respectively, in the controls, and 74.2% and 74.4%, respectively in the patients. As in Experiment 1, the proportion of studied items given a correct source judgment did not reliably differ for spatial versus temporal contexts, in neither the controls [t(13) = 1.3, p = 0.2] nor in the patients [t(6) = 1.7, p = 0.13]. Given the greater tendency for incorrect source judgments for the temporal than spatial context in the patients [t(6) = 6.6, p = 0.001] and controls [t(13) = 4.5, p = 0.001], source accuracy was also calculated using Pr as in Experiment 1. These estimates are shown in Figure 4.

Given previous evidence suggesting disproportionate impairments for source relative to item memory in frontal lesion patients, we conducted a Context (spatial, temporal) X Memory (item, source) X Group (controls, patients) ANOVA on the Pr measures. Given our a priori prediction for a greater impairment on temporal than spatial source memory for patients relative to controls, we cite one-tailed p-values for these interactions. The ANOVA revealed reliable main effects of Context [F(1, 19) = 28.1, p < 0.0001], Memory [F(1, 19) = 113.2, p < 0.0001] as well as a Context X Memory X Group interaction [F(1, 19) = 2.5, p < 0.05]. A follow-up Context X Group ANOVA for item memory revealed no main effects or interactions [F(1, 19)’s < 1]. In contrast, the follow-up ANOVA for source memory estimates revealed a reliable interaction [F(1, 19) = 2.6, p < 0.05]. Pairwise contrasts showed that there was no significant difference between spatial and temporal item memory within either group, and that neither of these estimates differed between controls and patients [t’s < 1]. In contrast, for both groups, source memory accuracy was greater for spatial than temporal contexts: patients [t(6) = 4.3, p = 0.005], controls: [t(13) = 3.6, p = 0.003]. Furthermore, while both spatial and temporal source accuracy were significantly greater than chance (0%) in the controls [t(13)’s > 5.1, p’s < 0.001], spatial [t(6) = 4.0, p = 0.007], but not temporal [t(6) < 1], source accuracy exceeded chance in the patients. Finally, pairwise contrasts confirmed that temporal [t(19) = 2.5, p = 0.01] but not spatial [t(19) < 1] source memory accuracy was impaired in the patients relative to the controls, as can be seen in Figure 4.

For consistency with Experiment 1, an ANOVA employing factors of Context (spatial, temporal), Source Response (correct, don’t know) and Group (controls, patients) for item recognition RTs, shown in Table 4, yielded a main effect of Source Response [F(1, 19) = 14.5, p = 0.001]. Pairwise comparisons revealed that correct source responses were faster than don’t know source responses for both contexts and for both groups [p’s < 0.05]. The same ANOVA performed for the source memory judgment RTs, shown in Table 4, yielded significant Context X Source Response, Context X Group and Source Response X Group interactions [F(1, 19)’s > 6.8, p’s < 0.02]. While there were no significant differences between these RTs in the controls [p’s > 0.15], the ANOVA for patients revealed a reliable effect of Context and a Context X Source Response interaction [F(1, 6)’s > 6.4, p’s < 0.02]. Pairwise comparisons showed that correct source responses were faster than don’t know responses in patients for the temporal task only [t(6) = 2.6, p = 0.04]. Finally, an ANOVA performed for the study phase RTs for items subsequently associated with correct source and don’t know spatial and temporal source judgments, shown in Table 3, revealed a main effect of Source Response [F(1, 19) = 11.1, p = 0.004]. As can be seen in the table, response times were faster for items subsequently associated

---

2 No reliable interactions between subsequent source memory accuracy and context (spatial, temporal) were observed at study in the left gyrus rectus region [x = −18, y = 27, z = −15] implicated in a marginal interaction during test [maximum T = 0.59, minimum SVC-corrected p = 0.28]. Similarly, no reliable interactions between source memory accuracy and context were observed at test in the right medial orbitofrontal [x = 9, y = 54, z = −12] or right gyrus rectus [x = 6, y = 33, z = −24] regions implicated during study [maximum T = 0.74, minimum SVC-corrected p = 0.23]
with don’t know source responses than with correct source responses across tasks for both groups.

DISCUSSION

By combining both imaging and neuropsychological methods, this study provides novel evidence that the orbitofrontal cortex (OFC) contributes to successful encoding and, to a lesser extent, retrieval of associations between objects and their temporal context. Indeed, the patient data in Experiment 2 suggest that the OFC is necessary for above-chance temporal context memory performance. While previous studies have suggested that OFC lesions produce difficulty in determining the temporal sequence of events (Schnider, 2003; Schnider & Ptak, 1999), the current results go further in showing that OFC contributes to both the encoding and retrieval of temporal context to a greater extent than at least some types of non-temporal (i.e. spatial) context.

Although the patients’ lesions were centered in medial aspects of the OFC, at least one patient had additional damage to the lateral OFC, making it difficult to determine which specific regions were necessary. The imaging results provided more precise anatomical information, with activity related to temporal context memory found in bilateral OFC, primarily in the gyrus rectus and medial OFC. These fMRI results are consistent with a previous PET study reporting disproportionate involvement of gyrus rectus to retrieval of temporal order relative to spatial location (Fujii et al., 2004), though this study only examined retrieval, the present event-related fMRI analyses also extend these findings by showing that OFC activity relates specifically to successful (relative to unsuccessful) encoding and retrieval of temporal context. It should be noted, however, that the interaction between source memory accuracy and context was only marginally reliable at retrieval, although follow up analyses show reliable OFC activity associated with temporal but not spatial context memory during retrieval.

The specific deficits in temporal context memory in the present patient data are consistent with theories suggesting that OFC damage can produce “temporal context confusions” or confusions in the temporal sequence of otherwise intact memory traces (e.g. (Schnider, 2003; Schnider & Ptak, 1999). These confusions in temporal order have been proposed to account for the confabulations sometimes observed in OFC patients. Although the patients in the current study did not overtly confabulate, prior work has also shown that frontal patients, including those with focal OFC damage, can exhibit temporal context impairments without confabulation (Gilboa et al., 2006; M. K. Johnson, O’Connor, & Cantor, 1997; Kopelman et al., 1997; Schnider, von Daniken, & Gutbrod, 1996), so these phenomena need not coincide. Regardless, one important way in which the current results differ from previous findings is that, in these studies (Gilboa et al., 2006; Schnider, 2003; Schnider & Ptak, 1999), memory for temporal information was measured indirectly via a continuous recognition task in which stimuli were repeatedly presented both within and across runs. Patients indicated stimulus repetitions within runs and the failure to inhibit positive responses to repetitions from previous runs was taken as an indicator of impaired temporal labeling or monitoring. Given the nature of this task, it may be that patients’ impairments reflected an inability to inhibit irrelevant memories (i.e., proactive interference), rather than temporal context impairments per se. By contrast, the current task measured temporal, and spatial, context memory directly, providing more direct evidence of temporal context memory impairments following OFC damage. Although the OFC may play a role in resolving interference, similar to the lateral PFC (Nee, Wager, & Jonides, 2007), susceptibility to interference is most likely similar for the spatial and temporal tasks in the current study.

One alternative explanation for the disproportionate involvement of the OFC in temporal context memory is that our temporal retrieval task may have been more difficult than the spatial
retrieval task, thereby increasing demands on cognitive control processes and frontal involvement (Rajah, Ames, & D'Esposito, 2008). Although accuracy was significantly lower in the temporal than the spatial task in both experiments, the data are not consistent with this hypothesis. Firstly, increased difficulty would also be expected to increase reaction times. However, neither the participants in Experiment 1 nor the controls in Experiment 2 showed reliable differences in reaction times at test between the two types of context for either the item judgment or the source recognition judgments. Only the patients exhibited longer reaction times for temporal than spatial source recognition judgments, consistent with their greater temporal than spatial source memory deficits. Secondly, it is unclear how a difficulty account could explain the reliable fMRI differences between temporal and spatial contexts that we found at encoding, which was associated with an incidental task, and also not associated with any differences in RTs between subsequent temporal and spatial source judgment in either experiment. It should be noted that while activity was greater for successful relative to unsuccessful (“don't know”) temporal source judgments during study, the opposite was true during test (i.e. activity was greater for don't know judgments). Whatever the reason for this different direction of activity during successful encoding and retrieval of temporal context, these findings would also appear difficult to explain entirely in terms of difficulty. That is, previous studies have suggested that, across numerous tasks, difficulty typically modulates activity in the dorsolateral PFC (e.g. Barch et al., 1997; Klingberg, O'Sullivan, & Roland, 1997; Rajah et al., 2008), a region that was spared in the patients here. Although it remains possible that the OFC may also be sensitive to task difficulty, further work is necessary to elucidate such a relationship.

It is worth discussing the possible contribution of item memory strength to the present temporal context judgments. That is, it is possible that the judgments could be based on the perceived recency of the item's previous presentation, for example, a decaying "strength" signal, without recollection of its precise temporal context (Hintzman, 2005). It is possible therefore that the greater involvement of OFC in our task reflects a role for OFC in judging recency. While we cannot rule this out based on this data set, it is noteworthy that item recognition performance was not impaired in the patients. This was the case even when restricting item recognition to familiarity-based judgments (Supplemental Material). Given that one might expect judgments of recency to contribute to (familiarity-based) item recognition, the recency account of temporal context impairments in the patients seems unlikely. Furthermore, although some OFC regions exhibited activity associated with item recognition in the imaging experiment, this was equivalent for the temporal and spatial tasks and was located in the lateral aspects of the OFC, where temporal context memory effects were not present.

There are undoubtedly many brain areas and related processes that contribute to memory for temporal context and as such, the OFC is not the only brain region that has been implicated in memory for temporal order. Numerous studies have shown that patients with lateral PFC damage also exhibit impairments in memory for temporal order (Janowsky, Shimamura, & Squire, 1989; Kesner, Hopkins, & Fineman, 1994; Kopelman et al., 1997; Mangels, 1997; Milner et al., 1991; Shimamura et al., 1990), and in some cases, relatively preserved memory for spatial location (Kopelman et al., 1997). While it is possible that the lateral PFC also plays a disproportionate role in memory for temporal context, lesions in these previous studies were either not well-described or were large and encompassed additional PFC regions, including the OFC. Moreover, memory for non-temporal context was not always assessed. Furthermore, other studies have shown that lateral PFC damage can also impair memory for conceptual information, spatial and other perceptual details (e.g. gender of voice) (Duarte, Ranganath, & Knight, 2005; Janowsky et al., 1989; Owen, Sahakian, Semple, Polkey, & Robbins, 1995; Swick, Senkfor, & Van Petten, 2006). Indeed, we previously found that activity in the lateral PFC was associated with source memory accuracy across spatial and temporal retrieval tasks (Duarte, Henson et al., 2008). Given the various cognitive control functions (e.g. sustained...
attention, monitoring, cue specification) attributed to lateral PFC (reviewed in Duncan & Owen, 2000), it seems likely that this region may play a more general “executive” role, contributing to memory for various kinds of contexts. Moreover, in contrast to the lateral PFC patients in many of these previous studies, the OFC patients here did not demonstrate significant impairments in neuropsychological tests of executive function. In a similar vein, patients with MTL lesions, particularly in the hippocampus, can also exhibit temporal order impairments (Hopkins, Kesner, & Goldstein, 1995; Kopelman et al., 1997). Given the various contextual details that patients with hippocampal damage have difficulty remembering, however, it seems likely that the hippocampus may contribute more generally to recollection (J. D. Johnson & Rugg, 2007), albeit disproportionately for spatial information, given the association of this region and the MTL generally with spatial processing (Eichenbaum, Yonelinas, & Ranganath, 2007; Lee et al., 2005).

It is worth discussing how the present results might fit into existing models of OFC functioning. For example, imaging studies have implicated the OFC in the processing of emotional information (Adolphs, 2002) and in decision making dependent on intuitive “feelings of rightness” (Elliott, Dolan, & Frith, 2000). Consistent with this, patients with OFC damage demonstrate impaired accuracy in their metamemory judgments (Schnyer et al., 2004). Furthermore, patients with OFC damage often exhibit difficulty in emotional regulation and in altering behavior to meet changes in stimulus-reinforcer associations (Rolls, 2004), such as is necessary for successful performance in the gambling tasks for which these patients are impaired (Bechara, 2004). Remembering the temporal context associated with an event is likely a complex process dependent on multiple component mechanisms and brain areas contributing to both encoding and retrieval (M.K. Johnson, Hashtroudi, & Lindsay, 1993; Marshuetz & Smith, 2006). Neuroanatomical studies in nonhuman primates have revealed major bidirectional connections between the OFC and MTL, including the hippocampus, as well as with the lateral PFC (reviewed in Petrides, 2007). Moreover, the connections of the OFC with the amygdala and autonomic pathways may underlie the feeling of rightness decisions associated with the OFC (Elliott et al., 2000). The OFC is therefore in a unique position to monitor the internal state (i.e. emotion, feeling of rightness) while mediating between the binding of episodic information by the MTL and the monitoring and evaluation of this information by the lateral PFC (Petrides, 2007). Although the current study was not designed to investigate this hypothesis, it is possible that the encoding and subsequent retrieval of temporal context taps into such reflective processes to a greater extent than does the encoding and retrieval of spatial location. Given that posterior cortical regions, including the MTL and parietal cortex may be particularly sensitive to spatial information (Eichenbaum et al., 2007; Husain & Nachev, 2007; Spiers & Maguire, 2007), successful memory for temporal information may necessitate more reflective processing mediated by regions like the OFC. Such a hypothesis is consistent with the idea that impaired feelings of rightness may contribute to temporal context confusions in OFC lesion patients (Gilboa et al., 2006). Future work is necessary to elucidate the relationship between hypothesized roles of the OFC and memory for temporal information.

In conclusion, the present results demonstrate that the OFC makes important contributions to memory for temporal context. By utilizing both imaging and patient methodologies, we have shown specifically that the medial aspect of the OFC, in the gyrus rectus and medial orbitofrontal gyrus, is associated with successful encoding and subsequent retrieval of temporal context, and is necessary for above-chance temporal context memory. More generally, our findings are consistent with the idea that memory-related processing may differ according to the type of information contained in the episode, as well as the type of information one is trying to retrieve (Wheeler and Buckner, 2003; Hornberger et al., 2006; Johnson and Rugg, 2007; Taylor et al., 2007).
Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Clay Clayworth for lesion reconstructions and Donatella Scabini for patient recruitment. This research was supported by the Medical Research Council (WBSE U.1055.05.012.00001.01), the Alzheimer’s Research Trust and NINDS grants NS21135 and PO1NS40813

References

Benton, AL.; Hamsher, SKd; Sivan, AB. Multilingual aplasia examination. 2nd ed.. Iowa City: AJA Associates; 1983.


Johnson JD, Rugg MD. Recollection and the Reinstatement of Encoding-Related Cortical Activity. Cereb Cortex. 2007


Figure 1.
(A) AAL ROIs for the 5 orbitofrontal regions investigated in Experiment 1. (B) Lesion overlap for the orbitofrontal patients studied in Experiment 2. The scale indicates the percentage of patients with damage to a particular area.
Figure 2.
Individual patient lesion reconstructions for the orbitofrontal patients studied in Experiment 2.
Figure 3.
Orbitofrontal regions demonstrating greater temporal than spatial source memory effects in Experiment 1. Plots show parameter estimates for the event-related response at the peak maxima within the selected ROI for each of the trial types versus the “baseline” condition: subsequent miss (forgotten) trials for encoding and correct rejection trials for retrieval (units arbitrary). Error bars depict standard error of the mean across participants. Asterisks indicate significant simple effects ($p < 0.05$ FWE). Note that size of activations cannot be compared across encoding and retrieval due to the use of different baselines.
Figure 4.
Source memory accuracy for both spatial and temporal retrieval tasks for control and orbitofrontal patient groups in Experiment 2. Error bars represent the standard error of mean across participants.
Table 1

Neuropsychological test scores from Experiment 2.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Controls (n = 14)</th>
<th>Patients (n = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAVLT Delayed Recall (# correct)</td>
<td>12.2 (2.5)</td>
<td>11.3 (3.9)</td>
</tr>
<tr>
<td>RAVLT Delayed Recognition (# correct)</td>
<td>14.3 (0.9)</td>
<td>12.8 (3.2)</td>
</tr>
<tr>
<td>COWAT Verbal Fluency (#correct)</td>
<td>44.6 (8.7)</td>
<td>44.0 (11.5)</td>
</tr>
<tr>
<td>WCST (# errors)</td>
<td>15.5 (12.2)</td>
<td>14.4 (15.2)</td>
</tr>
<tr>
<td>WMS-R Digit Span Forward</td>
<td>7.1 (0.86)</td>
<td>7.3 (0.76)</td>
</tr>
<tr>
<td>WMS-R Digit Span Backward</td>
<td>5.4 (1.1)</td>
<td>5.1 (1.0)</td>
</tr>
<tr>
<td>Rey Complex Figure Delayed Recall</td>
<td>21.3 (3.9)</td>
<td>19.0 (7.5)</td>
</tr>
</tbody>
</table>

Note: Standard deviations in parentheses. All neuropsychological tests are reported as raw scores.
Table 2

Proportions and corresponding reaction times to studied and unstudied items at test as a function of correct old/new judgment and subsequent correct, incorrect or “don’t know” source judgments for young adults in Experiment 1.

<table>
<thead>
<tr>
<th>Response</th>
<th>Item RT</th>
<th>Source RT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Studied items (Spatial)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct source</td>
<td>0.40 (0.12)</td>
<td>1527 (285)</td>
</tr>
<tr>
<td>Incorrect source</td>
<td>0.19 (0.04)</td>
<td>1527 (261)</td>
</tr>
<tr>
<td>Don’t know source</td>
<td>0.25 (0.12)</td>
<td>3420 (643)</td>
</tr>
<tr>
<td>New (Missed)</td>
<td>0.16 (0.09)</td>
<td>1573 (283)</td>
</tr>
<tr>
<td><strong>Studied items (Temporal)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct source</td>
<td>0.39 (0.08)</td>
<td>1537 (277)</td>
</tr>
<tr>
<td>Incorrect source</td>
<td>0.26 (0.06)</td>
<td>1546 (276)</td>
</tr>
<tr>
<td>Don’t know source</td>
<td>0.19 (0.09)</td>
<td>3355 (834)</td>
</tr>
<tr>
<td>New (Missed)</td>
<td>0.16 (0.09)</td>
<td>1568 (247)</td>
</tr>
<tr>
<td><strong>New items</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New (CR)</td>
<td>0.88 (0.09)</td>
<td>1336 (205)</td>
</tr>
</tbody>
</table>

Note: Standard deviations in parentheses. CR = Correct rejections, RT = Reaction Time. Item and source RTs are for initial item recognition and source memory decisions, respectively.
## Table 3

Reaction times during study as a function of subsequent memory judgments.

<table>
<thead>
<tr>
<th>Response</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls</td>
<td>Patients</td>
</tr>
<tr>
<td><strong>Spatial</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct source</td>
<td>1133 (168)</td>
<td>1044 (132)</td>
</tr>
<tr>
<td>Incorrect source</td>
<td>1134 (199)</td>
<td>1089 (180)</td>
</tr>
<tr>
<td>Don’t know source</td>
<td>1086 (228)</td>
<td>988 (133)</td>
</tr>
<tr>
<td>New (Missed)</td>
<td>1082 (242)</td>
<td>999 (192)</td>
</tr>
<tr>
<td><strong>Temporal</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct source</td>
<td>1122 (177)</td>
<td>1067 (140)</td>
</tr>
<tr>
<td>Incorrect source</td>
<td>1073 (199)</td>
<td>1017 (119)</td>
</tr>
<tr>
<td>Don’t know source</td>
<td>1042 (198)</td>
<td>859 (259)</td>
</tr>
<tr>
<td>New (Missed)</td>
<td>1062 (243)</td>
<td>1015 (194)</td>
</tr>
</tbody>
</table>

Note: Standard deviations in parentheses.
### Table 4

Proportions and corresponding reactions times to studied and unstudied items at test as a function of correct old/new judgment and subsequent correct, incorrect or "don’t know" source judgments for controls and patients in Experiment 2.

<table>
<thead>
<tr>
<th>Response</th>
<th>Controls</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Item RT</td>
<td>Source RT</td>
</tr>
<tr>
<td>Studied items (Spatial)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct source</td>
<td>0.42 (0.14)</td>
<td>2843 (1145)</td>
</tr>
<tr>
<td>Incorrect source</td>
<td>0.16 (0.09)</td>
<td>3459 (1297)</td>
</tr>
<tr>
<td>Don’t know source</td>
<td>0.29 (0.16)</td>
<td>3837 (1680)</td>
</tr>
<tr>
<td>New (Missed)</td>
<td>0.13 (0.09)</td>
<td>2368 (926)</td>
</tr>
<tr>
<td>Studied items (Temporal)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct source</td>
<td>0.39 (0.11)</td>
<td>2553 (1056)</td>
</tr>
<tr>
<td>Incorrect source</td>
<td>0.23 (0.09)</td>
<td>2625 (1209)</td>
</tr>
<tr>
<td>Don’t know source</td>
<td>0.26 (0.17)</td>
<td>2791 (1282)</td>
</tr>
<tr>
<td>New (Missed)</td>
<td>0.12 (0.07)</td>
<td>2457 (742)</td>
</tr>
<tr>
<td>New items</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New (CR)</td>
<td>0.97 (0.03)</td>
<td>1681 (456)</td>
</tr>
</tbody>
</table>

Note: Standard deviations in parentheses. CR = Correct rejections, RT = Reaction Time. Item and source RTs are for initial item recognition and source memory decisions, respectively.