Cogn Affect Behav Neurosci (2013) 13:900–915
DOI 10.3758/s13415-013-0170-x

Role of the hippocampus and orbitofrontal cortex during the disambiguation of social cues in working memory

Robert S. Ross · Matthew L. LoPresti · Karin Schon · Chantal E. Stern

Published online: 3 May 2013

Abstract Human social interactions are complex behaviors requiring the concerted effort of multiple neural systems to track and monitor the individuals around us. Cognitively, adjusting our behavior on the basis of changing social cues such as facial expressions relies on working memory and the ability to disambiguate, or separate, the representations of overlapping stimuli resulting from viewing the same individual with different facial expressions. We conducted an fMRI experiment examining the brain regions contributing to the encoding, maintenance, and retrieval of overlapping identity information during working memory using a delayed match-to-sample task. In the overlapping condition, two faces from the same individual with different facial expressions were presented at sample. In the nonoverlapping condition, the two sample faces were from two different individuals with different expressions. fMRI activity was assessed by contrasting the overlapping and nonoverlapping conditions at sample, delay, and test. The lateral orbitofrontal cortex showed increased fMRI signal in the overlapping condition in all three phases of the delayed match-to-sample task and increased functional connectivity with the hippocampus when encoding overlapping stimuli. The hippocampus showed increased fMRI signal at test. These data suggest that lateral orbitofrontal cortex helps encode and maintain representations of overlapping stimuli in working memory, whereas the orbitofrontal cortex and hippocampus contribute to the successful retrieval of overlapping stimuli. We suggest that the lateral orbitofrontal cortex and hippocampus play a role in encoding, maintaining, and retrieving social cues, especially when multiple interactions with an individual need to be disambiguated in a rapidly changing social context in order to make appropriate social responses.

Keywords Prefrontal · Social interaction · fMRI · Delayed match-to-sample

R. S. Ross · M. L. LoPresti · K. Schon · C. E. Stern
Center for Memory and Brain, Boston University, Boston, MA 02215, USA

R. S. Ross · M. L. LoPresti · K. Schon · C. E. Stern
Center of Excellence for Learning in Education, Science, and Technology—CELEST, Boston University, Boston, MA 02215, USA

R. S. Ross · M. L. LoPresti · K. Schon · C. E. Stern
Department of Psychology, Boston University, Boston, MA 02215, USA

R. S. Ross · M. L. LoPresti · K. Schon · C. E. Stern
Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA 02129, USA

C. E. Stern
Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA 02129, USA

R. S. Ross (✉)
Boston University, 2 Cummington Mall, Room 109, Boston, MA 02215, USA
e-mail: bross@bu.edu

The ability to perceive, maintain, and distinguish between different instances of encountering an individual is critical from both memory and social cognition perspectives. For example, in addition to being able to recognize a friend and distinguish one friend from another, it is also important that we identify changing moods in individuals by separately encoding changing facial expressions over both short- and long-term social interactions. The hippocampus and orbitofrontal cortex may be critical to guiding appropriate real-world social behavior during social gatherings when we

need to monitor the changing facial expressions of an individual. Animal studies have suggested that the hippocampus disambiguates, or separates, overlapping sequences (Agster, Fortin, & Eichenbaum, 2002; Bower, Euston, & McNaughton, 2005; Ginther, Walsh, & Ramus, 2011; Wood, Dudchenko, Robitsek, & Eichenbaum, 2000), and neuroimaging studies have shown hippocampal activation when learning (Kumaran & Maguire, 2006; Shohamy & Wagner, 2008) and retrieving (Brown, Ross, Keller, Hasselmo, & Stern, 2010; Ross, Brown, & Stern, 2009) overlapping sequences. Additionally, neuroimaging studies (LoPresti et al., 2008; McIntosh, Grady, Haxby, Ungerleider, & Horwitz, 1996; Olsen et al., 2009; Ranganath, Cohen, & Brozinsky, 2005; Ranganath & D'Esposito, 2001; Schon et al., 2005; Schon, Hasselmo, Lopresti, Tricarico, & Stern, 2004; Schon, Ross, Hasselmo, & Stern, 2013; Stern, Sherman, Kirchhoff, & Hasselmo, 2001), studies in patients with medial temporal lobe damage (Hannula, Tranel, & Cohen, 2006; Hartley et al., 2007; Nichols, Kao, Verfaellie, & Gabrieli, 2006; Olson, Moore, Stark, & Chatterjee, 2006; Olson, Page, Moore, Chatterjee, & Verfaellie, 2006), and electrophysiological recordings (Axmacher, Elger, & Fell, 2009) have suggested that the hippocampus is involved during the maintenance of information in working memory. Together, the findings that the hippocampus is involved in working memory and during the disambiguation of overlapping stimuli in long-term memory suggest that the hippocampus may be involved when using social cues to disambiguate multiple presentations of the same person over short delay periods.

In addition to the hippocampus, the orbitofrontal cortex may also be a critical region when disambiguating two encounters with the same individual in working memory. The orbitofrontal cortex is central to theories of social reinforcement learning (Kringelbach & Rolls, 2003; Rolls, 2004, 2007; Tabbert, Stark, Kirsch, & Vaitl, 2005). Importantly, the maintenance of socially relevant information such as faces (Courtney, Ungerleider, Keil, & Haxby, 1996; Haxby, Petit, Ungerleider, & Courtney, 2000; LoPresti et al., 2008; Sala, Rama, & Courtney, 2003) and emotional expressions (LoPresti et al., 2008) over short delays activates the orbitofrontal cortex. Along with its role in processing social information, the orbitofrontal cortex has been linked to the disambiguation of overlapping representations in long-term memory. In general, lateral orbitofrontal cortex is more strongly functionally connected to the hippocampus when retrieving disambiguated overlapping sequences (Brown, Ross, Tobyne, & Stern, 2012; Ross, Sherrill, & Stern, 2011) from long-term memory. Additionally, the orbitofrontal cortex is critical for tasks in which interference caused by repeated stimulus exposures must be overcome, such as in reversal learning (Berlin, Rolls, & Kischka, 2004; Chudasama & Robbins, 2003; Fellows & Farah, 2003; Hornak et al., 2004; McAlonan & Brown, 2003; Meunier, Bachevalier, & Mishkin, 1997; Rudebeck & Murray, 2008; Schoenbaum, Setlow, Nugent, Saddoris, & Gallagher, 2003; Tsuchida, Doll, & Fellows, 2010), proactive interference (Caplan, McIntosh, & De Rosa, 2007), and delayed nonmatch-to-sample working memory tasks involving small stimulus sets (LoPresti et al., 2008; Otto & Eichenbaum, 1992; Schon, Tinaz, Somers, & Stern, 2008). Together, the findings that the orbitofrontal cortex is involved in mnemonic tasks in which interference is present and that the orbitofrontal cortex is important for processing social cues suggest that the orbitofrontal cortex may interact with the hippocampus to create, maintain, and retrieve overlapping representations of the same individual seen with multiple facial expressions in working memory.

To investigate the contributions of hippocampus and orbitofrontal cortex to the encoding, maintenance, and retrieval of overlapping identity information during working memory, participants performed a modified delayed match-to-sample working memory task with pairs of face stimuli. In the overlapping condition, participants viewed a pair of pictures of the same individual shown with two different facial expressions. Importantly, the nonoverlapping condition consisted of two different faces with two different facial expressions. After the sample phase, the participants remembered the pair of stimuli across a short delay period. At test, they indicated whether a test face matched one of the sample faces shown before the delay. By comparing the overlapping and nonoverlapping conditions, the brain responses active during the disambiguation of two overlapping representations of the same person encountered with two different facial expressions during working memory encoding, maintenance, and retrieval could be examined. We also conducted a functional connectivity analysis to determine whether the orbitofrontal cortex and hippocampus work together when disambiguating overlapping stimuli in working memory. The results of the study provide insight into the brain mechanisms responsible during social situations in which it is important to keep different encounters with the same individual separate.

Materials and method

Participants

A group of 18 healthy individuals (seven male, 11 female; mean age = 19.2 years, *SD* = 1.17 years) with no history of neurological or psychiatric illness were recruited from the Boston University population for this study. Their vision was either normal or corrected to normal. All participants were screened for MRI environment compatibility. Eligible individuals who agreed to participate gave signed informed

consent in accordance with the Human Research Committee of the Massachusetts General Hospital and the Institutional Review Board of Boston University. One participant was unable to complete the study due to fMRI scanner malfunction, and another participant was eliminated from the study due to excess motion during scanning, leaving 16 participants for analysis.

Procedure

Stimuli The stimuli for the task were selected from the University of Pennsylvania database of facial expressions (Gur et al., 2002) and from other databases (Ekman & Friesen, 1976; Lyons, Akamatsu, Kamachi, & Gyoba, 1998; Pantic, Valstar, Rademaker, & Maat, 2005). We selected a range of expressions from 120 individuals, for a total of 200 unique stimuli that appeared in only one trial of the task. All of the stimuli were cropped to 350 × 467 pixels at a resolution of 28.35 pixels/cm (12.35 cm × 16.47 cm), put on a gray background, and converted to grayscale. An oval mask was used to remove peripheral features (e.g., hair, clothes, and eye color) and to isolate the central facial features (see Fig. 1).

Stimuli preexposure Prior to scanning, participants were preexposed to neutral expressions from all 120 individuals used in the study in order to familiarize them with each face identity. These neutral expressions were not used in the scanning task. Preexposure consisted of presenting all 120 individual faces with neutral expressions three times while participants made a male/female judgment, a young/old judgment, and an attractive/not attractive judgment. The judgments were made in order to encourage participants to attend to the faces.

Delayed match-to-sample task Participants performed a modified delayed match-to-sample task during fMRI scanning (Fig. 1). Each trial consisted of a pair of sample faces presented sequentially for 2 s each, followed by an 8-s delay period, followed by a single test face presented for 2 s. A variable-length (8, 10, or 12 s) intertrial interval (ITI) separated each trial. The task consisted of two conditions that differed only in the type of faces presented during the sample phase. In the overlapping condition (OL), two images of the same individual were shown with two different facial expressions. The nonoverlapping (NOL) condition differed by presenting two individuals with two different expressions. During the test period, participants were instructed to press "1" if the test face matched the first sample face presented, press "2" if the test face matched the second sample face presented, or press "3" if the test face did not match either of the sample faces. The nonmatch trials contained test stimuli that were the same identity as one of the two sample faces, but with a different expression.

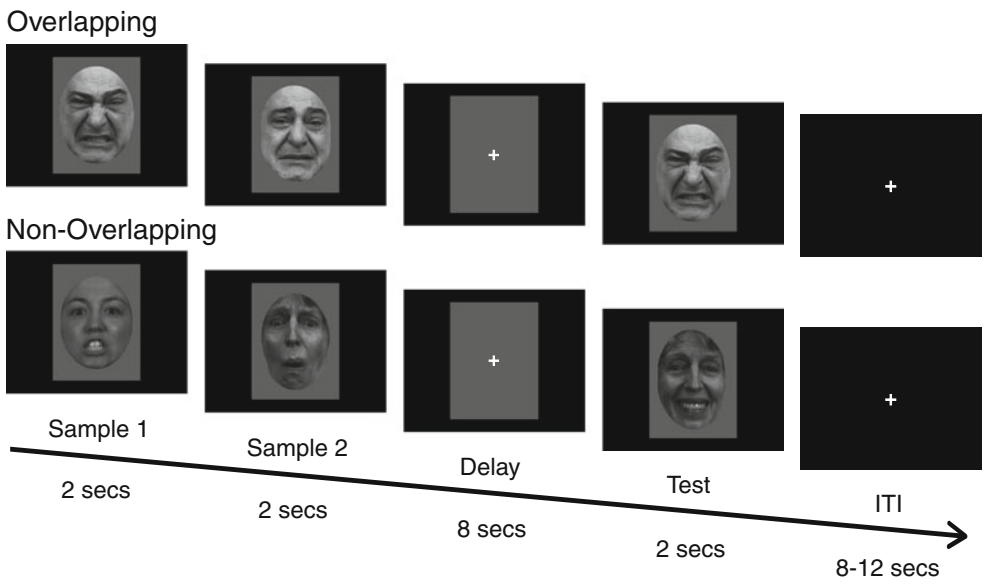

Fig. 1 Delayed match-to-sample task showing trials on which identity information was overlapping (OL) or nonoverlapping (NOL). A trial consisted of three time-locked components: a sample period in which two faces were presented sequentially for 2 s each, an 8-s delay period, and a 2-s test period in which a single face was presented. Trials were separated by 8-, 10-, or 12-s intertrial intervals (ITIs). During OL trials, participants were presented with a pair of sample faces from the same individual with different expressions. During NOL trials, participants were presented with a pair of sample faces from two different individuals with different expressions. In order for a trial to be a match, both the identity and the facial expression of the test face had to match one of the two sample faces. Nonmatch trials contained test stimuli that were the same identity as one of the two sample faces, but with a different expression. The OL trial shown in the top panels is an example of a match trial, and the NOL trial in the bottom panels is an example of a nonmatch trial

Overlapping/nonoverlapping conditions, match/nonmatch trials, and facial expressions were counterbalanced across five fMRI runs. Participants performed 16 trials per run, for a total of 80 trials (40 OL and 40 NOL). Participants viewed the task instructions and performed a practice version of the task during structural scanning. Responses and reaction times (RTs) were recorded from an MRI-compatible button box. The tasks were designed and presented and behavioral data were recorded with E-Prime 2 (Psychology Software Tools, Inc., Pittsburgh, PA).

fMRI data acquisition

Imaging was conducted on a 3.0-T Siemens MAGNETOM TrioTim scanner (Siemens AG, Medical Solutions, Erlanger, Germany) with a 12-channel Tim Matrix head coil at the Athinoula A. Martinos Center for Biomedical Imaging (Massachusetts General Hospital, Charlestown, MA). Two high-resolution T1-weighted multiplanar rapidly acquired gradient-echo (MP-RAGE) structural scans were acquired using generalized autocalibrating partially parallel acquisitions (GRAPPA) (TR = 2.530 s, TE = 3.44 ms, flip angle = 7°, slices = 176, field of view = 256 mm, resolution = 1 × 1 × 1 mm). Cognitive tasks were performed during functional T2*-weighted gradient-echo, echoplanar blood-oxygen-level-dependent (BOLD) scans (TR = 2 s, TE = 30 ms, flip angle = 90°, acquisition matrix = 64 × 64, field of view = 256 mm, slices = 32 interleaved axial–oblique, resolution = 4 × 4 × 4 mm, no interslice gap). Slices were aligned parallel to the line connecting the anterior and posterior commissures, and 192 images per run were acquired.

fMRI data analysis

Preprocessing Functional imaging data were preprocessed and statistically analyzed using the SPM8 software package (Statistical Parametric Mapping, Wellcome Department of Cognitive Neurology, London, UK). All BOLD images were first reoriented so that the origin (i.e., coordinate *xyz* = [0 0 0]) was at the anterior commissure. The images were then corrected for differences in slice timing and realigned to the first image collected within a series. Motion correction was conducted next and included realigning and unwarping the BOLD images to the first image in the series in order to correct for image distortions caused by susceptibility-by-movement interactions. Realignment was estimated using second-degree B-spline interpolation with no wrapping, whereas unwarp reslicing was done using fourth-degree B-spline interpolation with no wrapping. The high-resolution structural images were then coregistered to the mean BOLD image created during motion correction and were segmented into white- and gray-matter images. The bias-corrected structural images and the coregistered BOLD images were then spatially normalized into standard MNI (Montreal Neurological Institute) stereotactic space using the parameters derived during segmentation, with resampling of the BOLD images to 2 × 2 × 2 mm isotropic voxels. Finally, BOLD images were spatially smoothed using a 6-mm full-width-at-half-maximum Gaussian filter to reduce noise.

fMRI statistical analysis Analysis of fMRI activity during the delayed match-to-sample task was assessed with multiple regression using the SPM8 software package (the collinearity between the delay regressor and the sample and test regressors was ≤0.20). We used positive stick functions convolved with a Gamma hemodynamic response function (HRF; Boynton, Engel, Glover, & Heeger, 1996) in MATLAB 7.5 (The MathWorks, Inc., Natick, MA) to create 12 regressors that modeled the six components of the task (Sample 1, Sample 2, delay, test match, test nonmatch, and ITI) for each of the two conditions (overlapping and nonoverlapping). In addition, delay regressors were separated into four 1/4th-size stick functions spread across the four TRs (8 s) of the delay period to account for the sustained time course and expected weaker signal during this phase of the task (LoPresti et al., 2008; Schluppeck, Curtis, Glimcher, & Heeger, 2006). The five fMRI runs were concatenated in time and treated as a single times series. Additional regressors were included in the model to account for run number.

Linear contrasts were constructed to compare the overlapping condition to the nonoverlapping condition at the sample and delay periods of the task (i.e., OL sample > NOL sample, OL delay > NOL delay). Due to collinearity between the regressors for Samples 1 and 2, contrasts of the sample component consisted of a combination of both regressors. During the test period, participants were asked to identify the stimulus shown as a match or a nonmatch to the stimuli presented in the sample phase. At test, only the OL and NOL match trials were compared (OL match > NOL match). Group analysis was performed on each component of the task by entering the contrast images from each participant into a second-level random-effects one-sample *t* test, treating Participant as a random factor. Regions within the anterior medial temporal lobes (the amygdala, entorhinal cortex, and parts of perirhinal cortex) and anterior medial parts of the orbitofrontal cortex were not included because of signal dropout.

Functional connectivity analysis Functional connectivity was assessed using the beta series correlation analysis method (Rissman, Gazzaley, & D'Esposito, 2004). The beta series correlation method uses the magnitude of the task-related BOLD response for each individual trial to create a beta series. The method assumes that the extent to which two brain voxels interact during a given condition can be

quantified by the extent to which their respective condition-specific beta series are correlated. In order to conduct the beta series correlation analysis, we created a model in which each individual trial among the sample, delay, and test match trials was modeled separately for the overlapping and nonoverlapping conditions. Only correct trials were used in the model, so the number of regressors for each individual participant varied anywhere between 172 and 209 regressors. Additional regressors included as part of the design matrix were two regressors accounting for test nonmatch trials in the overlapping and nonoverlapping conditions, two regressors for the ITIs in the overlapping and nonoverlapping conditions, a nuisance regressor comprised of the sample, delay, and test onsets of incorrect trials, and five run regressors. All these regressors were then convolved with the canonical hemodynamic response function in SPM8 and filtered with a 0.008-Hz high-pass filter. In order to limit the amount of collinearity between the sample, delay, and test regressors, the first image in the sample was set as the onset for each sample regressor, whereas the onset point of the delay regressor was set to the second TR of the delay. Each regressor was modeled as a "stick" function (i.e., zero duration). Parameter estimates, or beta values, were computed for each regressor using the least-squares solution of the GLM in SPM8. We then sorted these beta values into the individual trials of the sample, delay, and test match periods of the overlapping and nonoverlapping conditions for the beta series correlation analysis. Beta series were formed within regions of interest (ROIs; see below) by concatenating the beta values for the individual trials of each condition (OL, NOL) and event (sample, delay, test) in chronological order.

We wished to determine regions showing differential functional connectivity with the orbitofrontal cortex in the overlapping as compared with the nonoverlapping condition at sample, delay, and test. We functionally defined the ROIs as 5-mm spheres centered at peak voxel coordinates in the orbitofrontal cortical regions identified in the univariate analysis. At sample, we examined functional connectivity with the right lateral orbitofrontal cortex (BA 47/12; MNI coordinate 44, 40, –10). At delay, functional connectivity was assessed with right lateral orbitofrontal cortex (BA 11L) at MNI coordinate 32, 42, –4. Two orbitofrontal cortical regions were examined for functional connectivity during test match trials, the right and left lateral orbitofrontal cortex (BA 47/12; MNI coordinates ±46, 34, 0). Eight correlation maps, one for each of the four ROIs for the overlapping and nonoverlapping conditions, were constructed by determining the correlation of each of the ROIs' beta series with the beta series of all other voxels in the brain using a custom MATLAB (The MathWorks, Natick, MA) script provided by Jesse Rissman. These correlation maps then underwent an arc-hyperbolic tangent transformation to normalize the values in the correlation maps. The arc-hyperbolic-transformed correlation coefficients were then divided by the standard deviation to produce a map of *z* scores. These *z*-score maps reflect how well each voxel in the brain is functionally connected to the orbitofrontal cortex ROI in the overlapping and nonoverlapping conditions separately.

The primary goal of the functional connectivity analysis was to determine whether the lateral orbitofrontal cortex was more strongly functionally connected to the hippocampus, caudate, and putamen when disambiguating overlapping social cues (overlapping vs. nonoverlapping condition). Therefore, we used the Volumes toolbox for SPM (http://sourceforge.net/projects/spmtools/) to extract the correlation *z* scores from each of the eight orbitofrontal cortex correlation maps (sample, delay, and two test coordinates, for the overlapping and nonoverlapping conditions) from the hippocampus (±30, –24, –15), caudate (±12, 22, –6), and putamen (±20, 18, –6) bilaterally in 5-mm-radius spheres. We used individual analyses of variance (ANOVAs) on the extracted correlation *z* scores for the sample, delay, and test periods to assess statistical significance. The hippocampal, caudate, and putamen coordinates used for the analysis were derived from two articles illustrating orbitofrontal–hippocampal and orbitofrontal–striatal functional connectivity during the disambiguation of long-term memories (Brown et al., 2012; Ross et al., 2011). A 2 × 2 × 3 repeated measures ANOVA of Condition (overlapping and nonoverlapping), Hemisphere (right and left), and Region (hippocampus, caudate, and putamen) was conducted on the correlation *z* values extracted from the orbitofrontal cortex connectivity maps at sample and delay separately. At test, two separate 2 × 2 × 3 ANOVAs were run, one for each orbitofrontal cortical correlation map (right and left lateral orbitofrontal cortex), with Condition, Hemisphere, and Region as factors. We also ran a group-level analysis to assess whole-brain differences in orbitofrontal cortical functional connectivity between the overlapping and nonoverlapping conditions. The whole-brain analysis was accomplished by conducting paired-sample *t* tests contrasting overlapping with nonoverlapping functional connectivity within SPM8 using the *z*-transformed correlation maps of orbitofrontal cortex connectivity at sample, delay, and test match trials.

A cluster extent threshold was enforced in order to correct for multiple comparisons for both the univariate and whole-brain functional connectivity analyses. Specifically, an individual-voxel statistical threshold of $p \leq .01$ was enforced with a minimum cluster extent threshold of 88 voxels (704 mm^3) in order to correct for multiple comparisons at $p \leq .05$. Therefore, at a voxel threshold of $p \leq .01$, the probability of observing a cluster extent larger than 88 voxels was $p \leq .05$. The cluster extent was calculated using a Monte Carlo simulation with 10,000 iterations run in MATLAB (Slotnick, Moo, Segal, & Hart, 2003). The Monte

Carlo simulation modeled activity in each voxel using a normally distributed random number (mean = 0 and variance = 1). Type I error was assumed to be equal to the individual-voxel threshold p value ($p \leq .01$) in a volume defined by the functional acquisition dimensions (64 × 64 × 32, with 4-mm isotropic original voxels resampled to 2-mm isotropic voxels with no masking). Spatial autocorrelation in the data was calculated for each participant and averaged about 7.5 mm after smoothing. Therefore, we used an 8-mm full-width-at-half-maximum three-dimensional Gaussian kernel in the Monte Carlo simulation.

Peaks within each cluster of activation after multiple-comparison correction were identified in SPM8. Peaks of activation were reported if they were more than 4 mm apart and represented a different region of activity. If a specific region of activity had multiple peaks within a cluster, the peak with the highest t value was reported. Brodmann areas were identified visually using a variety of reference materials (Damasio, 2005; Ongur, Ferry, & Price, 2003; Petrides, 2005, Scheperjans et al., 2008).

Behavioral analysis

Accuracy and RTs were recorded for each trial in E-Prime 2.0. A 2 × 2 repeated measures ANOVA of condition (overlapping and nonoverlapping) and trial type (match and nonmatch) was used to assess differences in accuracy and RTs individually. Pairwise comparisons and paired-sample t tests were used as post-hoc tests, where appropriate. The alpha level was set to $p \leq .05$. All ANOVAs and post-hoc tests were conducted using PASW 18, version 18.0.0 (IBM Corporation, NY). When more than four post-hoc tests were conducted for an individual ANOVA, the p values were Bonferroni-adjusted within PASW 18.

Results

Behavioral performance

The 2 × 2 repeated measures ANOVA examining accuracy revealed a significant main effect of condition [$F(1, 15) = 9.060$, $p = .009$], in which participants did significantly better on nonoverlapping (mean ± *SEM*: 93.0 ± 1.0 %) than on overlapping (88.8 ± 1.7 %) trials. However, we found no main effect of Trial Type (match vs. nonmatch), nor was there a Condition × Trial Type interaction. Some of the errors made in the overlapping condition were errors in which participants correctly identified the test stimulus as a match but incorrectly indicated the temporal order (i.e., the participant indicated that the test stimulus matched the first instead of the second sample, or vice versa). The number of these types of errors was small (3.7 %, or two trials per participant), precluding the possibility of using them as a condition for comparison. A 2 × 2 repeated measures ANOVA of Condition (overlapping or nonoverlapping) and Stimulus Order (first or second stimulus shown at sample) illustrated no effect of stimulus order in performance on match trials [$F(1, 15) = 4.099$, $p > .05$]. However, a significant Condition × Stimulus Order interaction did emerge [$F(1, 15) = 5.196$, $p \leq .05$]. The interaction was caused by a significant decrease in performance when the test stimulus in the overlapping trials was a match for the first stimulus shown at sample (81.9 ± 3.6 %), as compared with when the test stimulus matched the second stimulus (90.9 ± 2.3 %). The nonoverlapping condition showed no difference in performance on the basis of which sample picture was the match stimulus (first sample, 93.1 ± 2.2; second sample, 93.8 ± 1.4). The pattern of this interaction effect suggests that seeing a second stimulus with the same identity but a different facial expression causes interference that then needs to be resolved. We also observed a main effect of Condition for RTs [$F(1, 15) = 52.230$, $p \leq .001$], as well as a Condition × Trial Type interaction [$F(1, 15) = 79.451$, $p \leq .001$]. The main effect of Condition and the Condition × Trial Type interaction were driven by significantly faster RTs for match trials in the nonoverlapping condition than in the three other conditions, as assessed by paired-sample t tests [OL match, $t(15) = 10.811$, $p \leq .001$; OL nonmatch, $t(15) = 3.862$, $p \leq .01$; and NOL nonmatch, $t(15) = 3.543$, $p \leq .01$; see Fig. 2].

fMRI univariate analysis results

Encoding of overlapping stimuli in working memory Brain regions responsible for encoding overlapping representations of the same individual shown with two different facial expressions during working memory were identified by contrasting the OL and NOL sample regressors (Sample 1 and Sample 2) and revealed significant activation in the right lateral orbitofrontal cortex (OFC; xyz = 44, 40, –10; Fig. 3a). No other brain region showed a significant difference in activation after correcting for multiple comparisons (Table 1).

Maintenance of overlapping stimuli in working memory The brain regions responsible for maintaining overlapping representations of the same individual shown with two different facial expressions across a delay during working memory were identified by contrasting the OL and NOL delay regressors. This analysis revealed significant activation in right lateral orbitofrontal cortex (xyz = 32, 42, –4), in what Ongur et al. (2003) defined as Brodmann area (BA) 11L (Fig. 3b). Additionally, the left dorsal striatum, including both the putamen (xyz = −22, 10, –8) and the caudate (xyz = −16, 20, 0), as

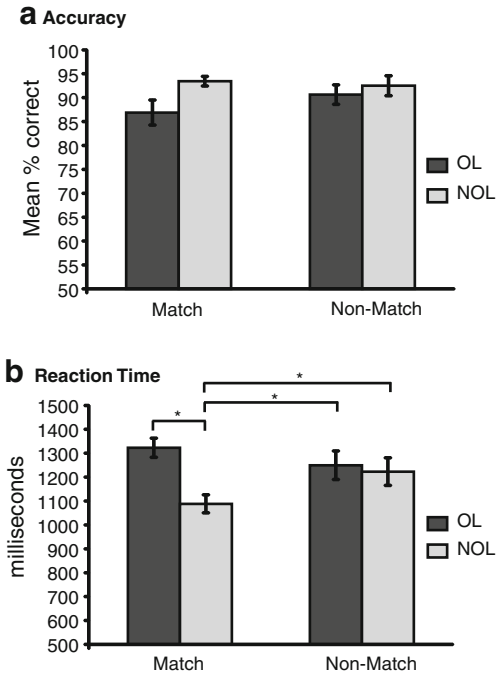

Fig. 2 Behavioral performance during the delayed match-to-sample task. (**a**) Mean percentages correct for overlapping (OL; dark gray bars) and nonoverlapping (NOL; light gray bars) conditions during match and nonmatch trials. (**b**) Reaction times (in milliseconds) for overlapping and nonoverlapping conditions during match and nonmatch trials. Error bars show standard errors of the means. *Significant differences between comparisons; the significance level was set at $p \leq .05$

well as the inferior temporal gyrus (IT; *xyz* = −52, –50, –8), showed increased fMRI activity in the delay period of the overlapping relative to the nonoverlapping condition (Table 1).

Retrieval of overlapping stimuli in working memory Brain regions active during the successful retrieval of an overlapping representation of an individual seen with two different facial expressions were determined by contrasting overlapping match trials with nonoverlapping match trials (Table 2). Significant fMRI activity was found in multiple parts of the lateral orbitofrontal cortex, including bilateral activity in Brodmann area 11L (left hemisphere *xyz* = −30, –32, –18; right hemisphere *xyz* = 30, 34, –14) and in Brodmann area 47/12 (left hemisphere *xyz* = −54, 34, –2; right hemisphere *xyz* = 46, 34, 0). Additionally, the posterior hippocampus showed bilateral activation (left hemisphere *xyz* = −32, –36, –8; right hemisphere *xyz* = 22, –40, –2) when participants successfully retrieved an overlapping stimulus (OL match vs. NOL match; Fig. 3c).

Comparison of nonoverlapping and overlapping conditions We directly contrasted the nonoverlapping condition to the overlapping condition at sample, delay, and test to examine brain regions showing more activation when encoding, maintaining, and retrieving two different faces with two different emotional expressions, as compared with when the same face was shown twice with different emotional expressions. Significant activations were apparent when comparing the nonoverlapping to the overlapping condition in the sample period of the task, including primary visual cortex (left hemisphere *xyz* = 14, –90, 2; right hemisphere *xyz* = −14, –92, 4), as well as left superior frontal gyrus (*xyz* = −16, 56, 20) and left anterior cingulate cortex (*xyz* = −8, 52, 14; Table 1). At delay and for test match trials, no brain regions showed more fMRI activity in the nonoverlapping than in the overlapping condition.

Functional connectivity results

ROI results

Orbitofrontal–hippocampal functional connectivity during encoding of overlapping stimuli The results of a 2 × 2 × 3 repeated measures ANOVA of Condition (overlapping and nonoverlapping), Hemisphere (right and left), and Region (hippocampus, caudate, and putamen) on the extracted correlational-map *z* scores at sample showed a significant main effect of Region [$F(2, 30) = 8.154$, $p \leq .001$] and a significant Condition × Region interaction [$F(2, 30) = 4.330$, $p \leq .05$]. Pairwise comparisons relating to the main effect of Region showed that collapsing condition and hemisphere, the orbitofrontal cortex was more strongly functionally connected to the hippocampus than to the caudate (OFC–hippocampus $z = 2.109 \pm 0.39$ [mean ± *SEM*] vs. OFC–caudate $z = 1.327 \pm 0.28$, $p \leq .05$) and was more strongly functionally connected to the putamen than the caudate (OFC–putamen $z = 2.292 \pm 0.35$ vs. OFC–caudate $z = 1.327 \pm 0.28$, $p \leq .05$). To further explore the significant Condition × Region interaction, three paired-sample *t* tests were run, directly comparing functional connectivity in the overlapping and nonoverlapping conditions in the hippocampus, caudate, and putamen, collapsing across hemispheres. Only three paired-sample *t* tests were run because we were interested in directly comparing the overlapping and nonoverlapping conditions within the same brain region. The results of these paired-sample *t* tests showed a significantly higher *z* score only for orbitofrontal connectivity to the hippocampus during the overlapping condition as compared to the nonoverlapping condition [$t(15) = 2.2$, $p \leq .05$; OFC–hippocampus: overlapping $z = 2.35 \pm 0.32$, nonoverlapping $z = 1.86 \pm 0.39$]. The higher *z* score for the overlapping than for the nonoverlapping condition suggests that when encoding overlapping stimuli, activity in the orbitofrontal cortex and hippocampus is more tightly correlated.

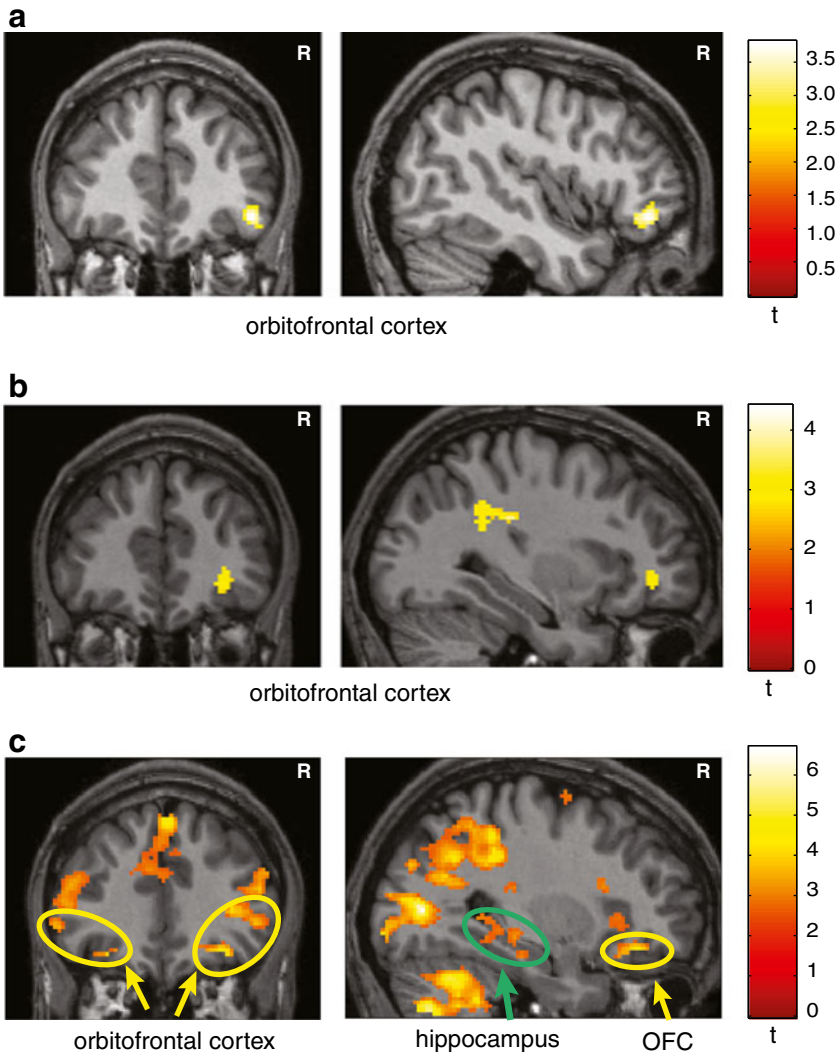

Fig. 3 Statistical parametric maps showing significantly greater fMRI activity during the disambiguation of overlapping stimuli in working memory. (**a**) fMRI activation in the orbitofrontal cortex ($y = 40$) related to the encoding of overlapping stimuli (OL sample > NOL sample). (**b**) fMRI activity in the orbitofrontal cortex ($y = 42$) during the maintenance of overlapping stimuli across a short delay period (OL delay > NOL delay). (**c**) Activations in orbitofrontal cortex bilaterally (OFC; lighter circles, yellow in online version) and the right hippocampus (darker circles, green online) during the successful retrieval of overlapping stimuli in working memory (OL match > NOL match). The displayed slices are from $y = 34$. For display purposes, the statistical parametric map is shown superimposed on a single participant's anatomical image ($p = .01$ with 88 contiguous voxels). R = right hemisphere

Orbitofrontal cortical functional connectivity differences during test match trials Two individual 2 × 2 × 3 repeated measures ANOVAs, one apiece for the right and left lateral orbitofrontal cortical correlation z scores, with Condition, Hemisphere, and Region (hippocampus, caudate, and putamen) as factors were conducted in order to assess functional connectivity differences during match trials at test. The 2 × 2 × 3 ANOVA examining z-score differences in the right lateral orbitofrontal cortex showed a significant main effect of region [$F(2, 30) = 15.208$, $p \leq .001$] and a significant Condition × Region interaction [$F(2, 30) = 4.624$, $p \leq .05$]. The main effect of region was caused by significantly higher z scores for connectivity between the right lateral orbitofrontal cortex and the putamen, relative to both the hippocampus (right lateral OFC–putamen $z = 1.71 \pm 0.22$ vs. right lateral OFC–hippocampus $z = 1.15 \pm 0.15$, $p \leq .01$) and the caudate (right lateral OFC–putamen $z = 1.71 \pm 0.22$ vs. right lateral OFC–caudate $z = 0.74 \pm 0.22$, $p \leq .001$), collapsing across conditions and hemispheres. Three paired-sample t tests were run in order to directly compare correlation z scores between the overlapping and nonoverlapping conditions within the hippocampus, caudate, and putamen, collapsing across hemispheres. Only the caudate–right lateral OFC connectivity showed a significant difference between the overlapping ($z = 0.4829 \pm 0.27$) and the nonoverlapping ($z = 0.9989 \pm 0.22$) conditions [$t(15) = 2.143$, $p \leq .05$]. The greater correlation z score for the nonoverlapping condition suggests that the right lateral OFC was more tightly correlated with caudate functioning during nonoverlapping test match trials.

We also found a significant main effect of Region in the left lateral orbitofrontal cortex [$F(2, 30) = 9.977$, $p \leq .001$] during the test period, although no significant Condition × Region interaction was found. As with the right lateral orbitofrontal cortex, the left lateral orbitofrontal cortex showed more functional connectivity with the putamen ($z = 1.42 \pm 0.18$) than with either the hippocampus

Table 1 Activity during encoding and maintenance of overlapping stimuli

Brain Region	k_E	Side	t	Coordinate	z
Overlapping sample > nonoverlapping sample					
orbitofrontal cortex (BA 47/12)	90	R	3.80	44, 40, –10	3.16
Nonoverlapping sample > overlapping sample					
superior frontal gyrus (BA 8)	808	L	5.19	−14, 44, 48	3.92
superior frontal gyrus (BA 9)		L	3.98	−16, 56, 20	3.27
anterior cingulate cortex (BA 32)		L	3.97	−8, 52, 14	3.27
calcarine sulcus (BA 17)	260	L	4.01	−16, –98, –4	3.29
calcarine sulcus (BA 17)	249	R	3.95	14, –90, 2	3.25
precuneus (7p)	148	L	3.22	−6, –60, 26	2.78
Overlapping delay > nonoverlapping delay					
putamen	115	L	4.08	−22, 10, –8	3.33
caudate		L	3.48	−16, 20, 0	2.96
inferior temporal gyrus (BA 21)	142	R	3.75	−52, –50, –8	3.13
orbitofrontal cortex (BA 11L)	110	R	3.13	32, 42, –4	2.72

k_E = cluster size in voxels. Regions active within the same cluster are grouped together. L, left; R, right; BA = Brodmann area. The t and z values are for the peak voxel. Coordinates are in MNI space

($z = 1.04 \pm 0.23$) or the caudate ($z = 0.60 \pm 0.18$), collapsing across conditions and hemispheres.

Whole-brain functional connectivity results

We examined whole-brain functional connectivity differences by directly contrasting the z maps from the beta-series correlation analysis corresponding to the overlapping and non-overlapping conditions. We used the same $p < .01$ voxel threshold, corrected for multiple comparisons, with a cluster extent of 88 voxels that had been used in the univariate analysis. The results of these contrasts can be viewed in Tables 3 and 4.

At sample, the right lateral orbitofrontal cortex (44, 40, –10) showed stronger functional connectivity with the posterior cingulate cortex during the overlapping than during the nonoverlapping condition (Table 3). Other significant differences in orbitofrontal cortical functional connectivity were in the comparisons of the nonoverlapping to the overlapping condition at delay and test. During the delay, the right lateral orbitofrontal cortex (32, 42, –4) showed stronger functional connectivity with the posterior cingulate cortex, parahippocampal cortex, and postcentral gyrus in the nonoverlapping condition as compared with the overlapping condition (Table 3).

Differential functional connectivity with the right and left lateral OFC (±46, 34, 0) during nonoverlapping relative to overlapping match trials at test can be seen in Table 4. Among other regions, the left lateral OFC showed stronger functional connectivity with the hippocampus bilaterally, the dorsal striatum bilaterally, and the left retrosplenial cortex during nonoverlapping match trials at test. The right lateral orbitofrontal cortex was also more strongly connected to the hippocampus bilaterally, as well as to the ventral lateral prefrontal cortex (BA 44) and to angular gyrus (BA 39) bilaterally during nonoverlapping match trials at test (Table 4).

Discussion

We contrasted fMRI activity and orbitofrontal cortical functional connectivity while participants encoded, maintained, and retrieved representations of the same individual shown with two different facial expressions (overlapping condition) with fMRI activity and orbitofrontal cortical functional connectivity while participants encoded, maintained, and retrieved pictures of two different individuals with two different facial expressions (nonoverlapping condition) in a delayed match-to-sample task. We observed significant right lateral orbitofrontal cortex (BA 47/12) activity during the sample phase and right lateral orbitofrontal cortex (BA 11L) activity during the delay phase when participants had to encode overlapping face stimuli and maintain disambiguated face representations across a short delay (Fig. 3a and b). Stronger functional connectivity was also apparent between the right lateral orbitofrontal cortex and the hippocampus at sample when encoding overlapping as compared to nonoverlapping face stimuli (Fig. 4). Additionally, during the test phase both the

orbitofrontal cortex and hippocampus showed greater fMRI activity in the overlapping than in the nonoverlapping condition (Fig. 3c).

Interestingly, though the hippocampus and orbitofrontal cortex showed increased fMRI activity for match trials at test in the overlapping condition, the lateral orbitofrontal cortex was more strongly functionally connected to the hippocampus, parahippocampal cortex, fusiform gyrus, and striatum during the nonoverlapping as compared to the overlapping test match trials (Table 4). In the nonoverlapping condition, two different individuals were shown at sample, each with a different facial expression. At test, the participants needed to determine which of the two sample identities the test face matched, suggesting that the increased orbitofrontal functional connectivity for nonoverlapping test match trials was due to retrieving two

Table 2 Activity when retrieving overlapping stimuli in working memory

Brain Region	k_E	Side	*t*	Coordinate	*z*
Overlapping match > nonoverlapping match trials					
fusiform gyrus (BA 37)	27,972	L	6.68	−34, –42, –14	4.55
fusiform gyrus (BA 37)		R	5.38	42, –42, –18	4.01
orbitofrontal cortex (BA 11L)		L	3.48	−30, 32, –18	2.96
orbitofrontal cortex (BA 47/12)		L	3.49	−54, 34, –2	2.96
hippocampus		R	5.19	22, –40, –2	3.92
hippocampus		L	5.27	−32, –36, –8	3.95
dorsal striatum		R	3.75	10, 8, 0	3.13
dorsal striatum		L	3.32	−8, 6, 0	2.85
calcarine fissure (BA 17)		R	6.43	30, –74, 6	4.46
intraparietal sulcus		L	5.37	−26, –60, 34	4.00
intraparietal sulcus		R	3.93	26, –60, 34	3.25
anterior cingulate cortex (BA 24)		L	5.00	−2, 28, 46	3.82
anterior cingulate cortex (BA 24)		R	4.74	10, 18, 40	3.69
ventral lateral prefrontal cortex (BA 45)		L	4.60	−46, 18, 20	3.62
dorsal lateral prefrontal cortex (BA 9/46)		L	4.81	−50, 22, 28	3.73
superior frontal gyrus (BA 8)		R	4.89	2, 32, 54	3.77
superior frontal gyrus (BA 8)		L	4.24	−4, 22, 60	3.42
middle frontal gyrus (BA 6)		L	4.18	−36, 0, 42	3.39
lingual gyrus (BA 18)		R	3.20	16, –72, 6	2.77
thalamus		R	5.90	10, –24, 12	4.24
thalamus		L	4.59	−6, –30, 0	3.61
anterior insula		L	5.62	−36, 20, –10	4.12
inferior temporal gyrus (BA 19)		R	5.54	46, –60, –6	4.08
lateral occipital gyrus (BA 18)		R	5.25	24, –90, –4	3.95
cerebellum		R	5.75	28, –72, –46	4.17
cerebellum		R	5.68	10, –72, –24	4.14
cerebellum		R	5.56	14, –50, –18	4.09
cerebellum		L	3.99	−6, –58, –20	3.28
cerebellum		L	6.39	−2, –70, –22	4.46
midbrain		L	5.24	−4, –24, –10	3.94
orbitofrontal cortex (BA 11L)	532	R	4.45	30, 34, –14	3.54
orbitofrontal cortex (BA 47/12)		R	3.65	46, 34, 0	3.08
anterior insula		R	4.11	46, 22, –12	3.35
dorsal lateral prefrontal cortex (BA 9/46)	807	R	5.59	50, 26, 26	4.11
middle frontal gyrus (BA 6)	132	L	5.10	−40, 2, 58	3.87
middle frontal gyrus (BA 6)	111	R	4.15	44, 0, 58	3.37

k_E = cluster size in voxels. Regions active within the same cluster are grouped together. L, left; R, right; BA = Brodmann area. The *t* and *z* values are for the peak voxel. Coordinates are in MNI space

Table 3 Whole brain functional connectivity with orbitofrontal cortex

Brain Region	k_E	Side	t	Coordinate	z
Overlapping sample > nonoverlapping sample (44, 40, –10)					
posterior cingulate cortex (BA 23)	481	L	5.00	−14, –60, 22	3.78
posterior cingulate cortex (BA 23)		R	4.67	12, –54, 24	3.61
Nonoverlapping delay > overlapping delay (32, 42, –4)					
posterior cingulate cortex (BA 31)	597	R	5.88	14, –18, 48	4.17
parahippocampal cortex	114	R	4.65	34, –40, –10	3.6
postcentral gyrus	323	R	3.56	36, –40, 62	2.99

k_E = cluster size in voxels. Regions active within the same cluster are grouped together. L, left; R, right; BA = Brodmann area. The t and z values are for the peak voxel. Coordinates are in MNI space

different identities, relative to only one identity in the overlapping condition.

These results have multiple implications for how social cues like face identity and emotional expression are encoded, maintained, and retrieved in working memory. When encountering the same familiar individual multiple times with different social cues, such as facial expression, the right lateral orbitofrontal cortex and hippocampus are more strongly recruited and show stronger functional connectivity when disambiguating each encounter with that individual. Disambiguating each encounter with the same individual ensures that the correct social cue will be used to select the appropriate response and to guide behavior during subsequent encounters. The right lateral orbitofrontal cortex and striatum show increased recruitment during the delay period after encountering the same individual with two different facial expressions (Fig. 3b, Table 1), suggesting that these regions play a role in maintaining overlapping stimuli. When reencountering the same individual previously seen with two different social cues, the lateral orbitofrontal cortex, along with the hippocampus and fusiform gyrus bilaterally, shows increased fMRI activity (Fig. 3c, Table 2), suggesting that these regions contribute to the retrieval of the correct previous encounter with that individual.

Orbitofrontal cortex and the encoding, maintenance, and retrieval of overlapping stimuli

The orbitofrontal cortex may contribute to memory by working with the hippocampus to separate when the same or overlapping stimuli have been experienced. The hippocampus sends anatomical projections to the orbitofrontal cortex (Barbas & Blatt, 1995; Cavada, Company, Tejedor, Cruz-Rizzolo, & Reinoso-Suarez, 2000; Insausti & Munoz, 2001; Roberts et al., 2007), providing the anatomical framework for a concerted engagement of these two regions in memory processes. Two recent studies suggested that the hippocampus and orbitofrontal cortex work together during the disambiguation of overlapping sequences in long-term memory (Brown et al., 2012; Ross et al., 2011). In the present study, we showed increased orbitofrontal cortical activity during the sample, delay, and test phases of a working memory task in which participants were asked to disambiguate two overlapping social stimuli. Importantly, the present study also showed that the right lateral orbitofrontal cortex works with the hippocampus when encoding overlapping stimuli in a working memory task. These results suggest that the orbitofrontal cortex interacts with the hippocampus to create separate representations of overlapping stimuli during the encoding phase of a working memory task and that the orbitofrontal cortex contributes to the maintenance and retrieval of overlapping stimuli.

One way that the orbitofrontal cortex may contribute to separating, or disambiguating, overlapping stimuli is to assist the hippocampus in linking each stimulus exposure to the specific context in which it was experienced. The orbitofrontal cortex is more strongly activated (Brown et al., 2010) and shows stronger functional connectivity (Brown et al., 2012; Ross et al., 2011) at choice points when contextual information is needed to guide decision making. Damage to the orbitofrontal cortex impairs the ability of monkeys to correctly learn the specific reward value of a repeatedly presented stimulus in any singular trial, as well as causing response strategies to shift to a more probabilistic strategy (Walton, Behrens, Buckley, Rudebeck, & Rushworth, 2010). The inability to learn the specific reward value of a repeated stimulus in any one trial suggests an inability to disambiguate each encounter. Additionally, the orbitofrontal cortex is critical for making correct decisions in reversal-learning tasks (Berlin et al., 2004; Chudasama & Robbins, 2003; Fellows & Farah, 2003; Hornak et al., 2004; McAlonan & Brown, 2003; Meunier et al., 1997; Rudebeck & Murray, 2008; Schoenbaum et al., 2003; Tsuchida et al., 2010). After reversal, the reward contingencies change when the previously rewarded stimulus becomes unrewarded and the previously unrewarded stimulus becomes rewarded. Importantly, the earlier stimulus–outcome associations still exist

Table 4 Lateral orbitofrontal functional connectivity at test

Brain Region	k_E	Side	t	Coordinate	z
Nonoverlapping match > overlapping match test trials (−46, 34, 0)					
retrosplenial cortex	1,013	L	6.36	−28, –72, 6	4.37
superior temporal gyrus (BA 22)		L	4.79	−46, –40, 14	3.67
hippocampus		L	4.46	−36, –20, –18	3.50
hippocampus		L	4.32	−34, –34, –4	3.43
cuneus (BA 19)	382	L	5.28	−6, –90, 32	3.91
cuneus (BA 19)		R	4.24	6, –8, 34	3.38
parietal–occipital sulcus		R	4.16	4, –66, 24	3.34
cuneus (BA 18)		L	3.26	−6, –100, 10	2.79
paracentral gyrus	356	R	4.86	8, –30, 60	3.71
paracentral gyrus		L	3.88	−12, –36, 62	3.18
postcentral gyrus		L	3.70	−16, –34, 64	3.07
dorsal striatum	431	R	4.02	22, 18, 20	3.26
dorsal striatum	172	L	3.77	−14, 20, 16	3.11
hippocampus	133	R	3.96	38, –30, –10	3.23
insula	136	R	3.69	32, –12, 18	3.06
retrosplenial cortex	106	L	3.10	−10, –50, 4	2.68
Nonoverlapping match > overlapping match test trials (46, 34, 0)					
ventral lateral prefrontal cortex (BA 44)	1,623	L	6.93	−60, –2, 16	4.57
insula		L	6.27	−44, –14, 4	4.33
precentral gyrus		L	4.94	−58, –8, 24	3.75
superior temporal gyrus (BA 22)		L	4.93	−52, –24, 6	3.74
ventral lateral prefrontal cortex (BA 44)	1,544	R	5.08	64, –2, 18	3.81
insula		R	4.84	40, –18, 18	3.70
superior temporal gyrus (BA 22)		R	4.49	50, –20, 6	3.52
precentral gyrus		R	4.33	58, –2, 14	3.43
hippocampus	92	L	5.06	−32, –32, –6	3.80
hippocampus	100	R	3.50	36, –38, –2	2.95
angular gyrus (BA 39)	96	L	4.24	−58, –58, 22	3.39
angular gyrus (BA 39)	142	R	3.98	42, –52, 18	3.24
medial rostral prefrontal cortex (BA 10)	249	L	6.71	−20, 46, 2	4,49

k_E = cluster size in voxels. Regions active within the same cluster are grouped together. L, left; R, right; BA = Brodmann area. The t and z values are for the peak voxel. Coordinates are in MNI space

after the reversal, creating overlapping representations that need to be disambiguated. Both monkeys (Rudebeck & Murray, 2008) and humans (Tsuchida et al., 2010) with orbitofrontal cortex damage make more mistakes postreversal, because they are more likely to switch their response after a rewarded trial to the previous reward contingency. Switching back to the old response contingency after orbitofrontal cortex damage suggests an inability to disambiguate the current reward contingency from prior contingencies. Therefore, we suggest that a critical function of the orbitofrontal cortex during both long-term and working memory tasks is to assist in resolving interference caused by overlapping representations, by linking each encounter with a stimulus to the context in which it appeared.

In the present study, we showed increased fMRI activity in the right lateral orbitofrontal cortex and increased functional connectivity with the hippocampus when encoding two overlapping social stimuli. Specifically, when the participant was shown two pictures of the same individual with different social cues (facial expressions), the right lateral orbitofrontal cortex and hippocampus were more strongly functionally connected than when viewing two pictures of two different individuals with different social cues. Viewing the same individual with different facial expressions creates interference between the two encounters with the individual. In order to correctly indicate whether the test face matched the first or second stimulus viewed at sample in the overlapping condition, the participants would have needed

to separate the two social cues (e.g., happy vs. sad face). Our results show that the orbitofrontal cortex and hippocampus work together when encoding the specific instance of encountering the same individual. We suggest that the ability to separate different encounters with the same individual seen with different social cues allows for appropriate social interactions. For example, if a person is seen first as happy and then as sad, the appropriate social response would change. We propose that the orbitofrontal cortex and hippocampus are part of a working memory system, as well as a long-term memory system, that allows us to flexibly encode separate representations of an individual in the varying social contexts in which we encounter him or her.

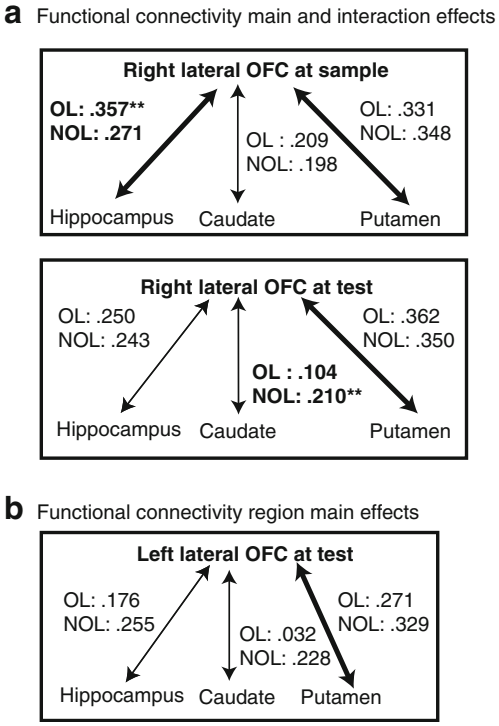

Fig. 4 Summary of functional connectivity between lateral orbitofrontal cortex (OFC) and the hippocampus, caudate, and putamen regions of interest. The *r* values of the correlation between the OFC and each region of interest for the overlapping (OL) and nonoverlapping (NOL) conditions are reported. (**a**) At sample (top), right lateral orbitofrontal–hippocampal functional connectivity at sample was stronger during the overlapping than during the nonoverlapping condition (indicated by **). Both the putamen and hippocampus were more strongly connected to the orbitofrontal cortex, collapsing across conditions (thickened arrows). At test (bottom), right lateral orbitofrontal–caudate functional connectivity was stronger during the nonoverlapping than during the overlapping condition (indicated by **), although the putamen showed the strongest functional connectivity with the orbitofrontal cortex when collapsing across conditions (thickened arrow). (**b**) The left lateral OFC was more strongly functionally connected to the putamen when collapsing across conditions (thickened arrow)

The hippocampus and the retrieval of overlapping stimuli in working memory

In order to correctly indicate whether the test stimulus matched the first or second stimulus shown in the sample period of the overlapping condition, participants needed to disambiguate overlapping representations (two different facial expressions) of a single individual. The hippocampus has been shown to be important for disambiguating, or separating, overlapping sequences (Agster et al., 2002; Kumaran & Maguire, 2006; Ross et al., 2009) and overlapping navigational routes (Brown et al., 2010; Wood et al., 2000) during long-term memory. Our results show that in the test phase of a delayed match-to-sample task, the hippocampus has stronger fMRI activity when retrieving a single face identity shown with two different facial expressions than when retrieving two different face identities shown with two different facial expressions. These data extend previous work by suggesting that the hippocampus contributes to working memory when retrieving representations of overlapping stimuli.

Contrary to our predictions, we did not see differential hippocampal fMRI activity during the encoding or maintenance of overlapping representations of the same individual. It may be that the present whole-brain scanning protocol lacked the necessary resolution to detect disambiguation-related hippocampal activity during encoding and maintenance. In support of this idea, recent work in our laboratory, using the same paradigm with high-resolution fMRI focused on the medial temporal lobes, has shown hippocampal activity differences at encoding and maintenance at the subfield level. Specifically, in that study we observed greater hippocampal activity in CA3/dentate gyrus and CA1 during encoding and in CA1 and the subiculum during the delay period of the overlapping condition, as compared with the nonoverlapping condition (Newmark, Schon, Ross, & Stern, 2013), suggesting a role for the hippocampus when encoding overlapping stimuli and when maintaining disambiguated stimuli in working memory.

It has recently been suggested that the hippocampus plays a role in working memory, although the nature of hippocampal involvement in working memory is under debate. Some researchers have suggested that the hippocampus is only involved when the capacity of working memory has been exceeded, resulting in the engagement of long-term memory processes in short-term memory tasks (Jeneson, Mauldin, & Squire, 2010; Shrager, Levy, Hopkins, & Squire, 2008). Other research has suggested that the hippocampus contributes to working memory in a process-specific manner. The hippocampus has been associated with long-term relational memory (Eichenbaum, 2000). In addition, neuropsychological experiments in human amnesic patients have shown that the hippocampus is also critical

to remembering relational information over short delay periods (Finke et al., 2008; Hannula et al., 2006; Nichols et al., 2006; Olson, Moore, et al., 2006; Olson, Page, et al., 2006). The present results showed increased hippocampal activity in the test phase of a working memory task in which disambiguating overlapping stimuli and maintaining disambiguated representations over a brief delay was critical to task performance. Together with prior studies showing hippocampal involvement in the disambiguation of long-term memories (Agster et al., 2002; Brown et al., 2010; Kumaran & Maguire, 2006; Ross et al., 2009; Wood et al., 2000) and a recent high-resolution fMRI study showing hippocampal activation during encoding and maintenance using the same task reported here (Newmark et al., 2013), our results provide evidence suggesting that disambiguation-related activity associated with hippocampal function is not dependent on the amount of time that the information is to be remembered, and is present in both long-term and working—that is, short-term—memory paradigms. This is critical for processing social cues, as they are dynamic and can change rapidly, requiring a mechanism by which individual encounters with an individual can be segregated so that appropriate behavioral responses can be selected in subsequent interactions.

Orbitofrontal cortical functional connectivity during test match trials

During match trials of the nonoverlapping condition, the lateral orbitofrontal cortex showed stronger functional connectivity with many brain regions, including the hippocampus, retrosplenial cortex, and fusiform gyrus, as compared with overlapping match trials. Unlike the overlapping condition, in which the same person was shown with two different facial expressions during the sample period, the nonoverlapping condition had two different individuals with two different facial expressions presented at sample. Therefore, the differences in orbitofrontal cortical functional connectivity at test between the nonoverlapping and overlapping conditions could be related to viewing two different identities at sample, suggesting that the increased functional connectivity may have been related to a load effect. The hippocampus and retrosplenial cortex have shown increased fMRI activity in the retrieval phase of a working memory task with increased load (Schon, Quiroz, Hasselmo, & Stern, 2009). In the present study, the hippocampus may have acted as a match/mismatch detector (Kumaran & Maquire, 2006, 2007) indicating that the test stimulus matched a previously viewed stimulus. The increased functional connectivity between the orbitofrontal cortex and the hippocampus, retrosplenial cortex, and fusiform gyrus may then have been caused by the participant retrieving two face identities from the sample in order to determine which one the test face matched.

The present study may not have shown increased fMRI activity in the nonoverlapping condition due to the interference inherent in viewing overlapping stimuli, during which more processing resources might be needed to determine which specific sample stimulus matched the test stimulus. In the overlapping condition, the hippocampus would have also indicated that a match was present, but due to the interference inherent in viewing overlapping stimuli, more processing resources would have been needed to determine which specific sample stimulus matched the test stimulus. The load effect for the nonoverlapping stimuli and the interference in the overlapping condition may explain why we found increased functional connectivity between the orbitofrontal cortex, hippocampus, retrosplenial cortex, and fusiform gyrus at test for the nonoverlapping condition, but increased fMRI activity during overlapping trials in these brain regions. It is important to note that the orbitofrontal cortical functional connectivity at test that we have reported here was the result of a contrast between the nonoverlapping and overlapping condition connectivity profiles. The results do not imply that there was no functional connectivity between the orbitofrontal cortex and these other brain regions in the overlapping condition, simply that the functional connectivity was stronger when matching a test stimulus in the nonoverlapping condition.

Conclusion

Combined with those from prior studies, our results suggest that the hippocampus and orbitofrontal cortex play roles in both working memory and long-term memory when separate representations of overlapping stimuli need to be disambiguated. In addition, we suggest that this ability to disambiguate overlapping representations is important for social interactions, in which the orbitofrontal cortex and hippocampus play key roles in enabling us to flexibly encode and distinguish between changing contexts and social situations, so that we may act appropriately.

Author note This work was supported by the National Science Foundation Science of Learning Center (Grant No. SMA-0835976), the National Institutes of Health (Grant No. P50 MH071702), and the Department of Psychology at Boston University. Functional magnetic resonance imaging was conducted at the Athinoula A. Martinos Center for Biomedical Imaging at the Massachusetts General Hospital, using resources provided by the Center for Functional Neuroimaging Technologies, Grant No. P41RR14075, a P41 Regional Resource supported by the Biomedical Technology Program of the National Center for Research Resources (NCRR), National Institutes of Health. This work also involved the use of instrumentation supported by the NCRR Shared Instrumentation Grant Program and the High-End Instrumentation Grant Program; specifically, Grant No. S10RR021110. We thank Matthew Grace for his assistance developing and piloting the behavioral task, and Michael Hasselmo and David Somers for helpful discussions about the study. We also thank Ruben Gur's Brain Behavior Laboratory at

the University of Pennsylvania for supplying the majority of the face stimuli used in this study.

References

Agster, K. L., Fortin, N. J., & Eichenbaum, H. (2002). The hippocampus and disambiguation of overlapping sequences. *Journal of Neuroscience, 22*, 5760–5768.

Axmacher, N., Elger, C. E., & Fell, J. (2009). Working memory-related hippocampal deactivation interferes with long-term memory formation. *Journal of Neuroscience, 29*, 1052–1960.

Barbas, H., & Blatt, G. J. (1995). Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. *Hippocampus, 5*, 511–533.

Berlin, H. A., Rolls, E. T., & Kischka, U. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain, 127*, 1108–1126.

Bower, M. R., Euston, D. R., & McNaughton, B. L. (2005). Sequential-context-dependent hippocampal activity is not necessary to learn sequences with repeated elements. *Journal of Neuroscience, 25*, 1313–1323.

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience, 16*, 4207–4221.

Brown, T. I., Ross, R. S., Keller, J. B., Hasselmo, M. E., & Stern, C. E. (2010). Which way was I going? Contextual retrieval supports the disambiguation of well learned overlapping navigational routes. *Journal of Neuroscience, 30*, 7414–7422.

Brown, T. I., Ross, R. S., Tobyne, S. M., & Stern, C. E. (2012). Cooperative interactions between hippocampal and striatal systems support flexible navigation. *NeuroImage, 60*, 1316–1330.

Caplan, J. B., McIntosh, A. R., & De Rosa, E. (2007). Two distinct functional networks for successful resolution of proactive interference. *Cerebral Cortex, 17*, 1650–1663.

Cavada, C., Company, T., Tejedor, J., Cruz-Rizzolo, R. J., & Reinoso-Suarez, F. (2000). The anatomical connections of the macaque monkey orbitofrontal cortex: A review. *Cerebral Cortex, 10*, 220–242.

Chudasama, Y., & Robbins, T. W. (2003). Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: Further evidence for the functional heterogeneity of the rodent frontal cortex. *Journal of Neuroscience, 23*, 8771–8780.

Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1996). Object and spatial visual working memory activate separate neural systems in human cortex. *Cerebral Cortex, 6*, 39–49.

Damasio, H. (2005). *Human brain anatomy in computerized images* (2nd ed.). New York, NY: Oxford University Press.

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews Neuroscience, 1*, 41–50.

Ekman, P., & Friesen, W. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Fellows, L. K., & Farah, M. J. (2003). Ventromedial frontal cortex mediates affective shifting in humans: Evidence from a reversal learning paradigm. *Brain, 126*, 1830–1837.

Finke, C., Braun, M., Ostendorf, F., Lehmann, T. N., Hoffmann, K. T., Kopp, U., & Ploner, C. J. (2008). The human hippocampal formation mediates short-term memory of colour–location associations. *Neuropsychologia*, *46*, 614–623.

Ginther, M. R., Walsh, D. F., & Ramus, S. J. (2011). Hippocampal neurons encode different episodes in an overlapping sequence of odors task. *Journal of Neuroscience, 31*, 2706–2711.

Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., . . . Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, *115*, 137–143.

Hannula, D. E., Tranel, D., & Cohen, N. J. (2006). The long and the short of it: Relational memory impairments in amnesia, even at short lags. *Journal of Neuroscience, 26*, 8352–8359.

Hartley, T., Bird, C. M., Chan, D., Cipolotti, L., Husain, M., Vargha-Khadem, F., & Burgess, N. (2007). The hippocampus is required for short-term topographical memory in humans. *Hippocampus*, *17*, 34–48.

Haxby, J. V., Petit, L., Ungerleider, L. G., & Courtney, S. M. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *NeuroImage, 11*, 380–391.

Hornak, J., O'Doherty, J., Bramham, J., Rolls, E. T., Morris, R. G., Bullock, P. R., & Polkey, C. E. (2004). Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience*, *16*, 463–478.

Insausti, R., & Munoz, M. (2001). Cortical projections of the non-entorhinal hippocampal formation in the cynomolgus monkey (*Macaca fascicularis*). *European Journal of Neuroscience, 14*, 435–451.

Jeneson, A., Mauldin, K. N., & Squire, L. R. (2010). Intact working memory for relational information after medial temporal lobe damage. *Journal of Neuroscience, 30*, 13624–13629.

Kringelbach, M. L., & Rolls, E. T. (2003). Neural correlates of rapid reversal learning in a simple model of human social interaction. *NeuroImage, 20*, 1371–1383.

Kumaran, D., & Maguire, E. A. (2006). The dynamics of hippocampal activation during encoding of overlapping sequences. *Neuron, 49*, 617–629.

Kumaran, D., & Maguire, E. A. (2007). Match–mismatch processes underlie human hippocampal responses to associative novelty. *Journal of Neuroscience, 27*, 8517–8524.

LoPresti, M. L., Schon, K., Tricarico, M. D., Swisher, J. D., Celone, K. A., & Stern, C. E. (2008). Working memory for social cues recruits orbitofrontal cortex and amygdala: A functional magnetic resonance imaging study of delayed matching to sample for emotional expressions. *Journal of Neuroscience, 28*, 3718–3728.

Lyons, M. J., Akamatsu, S., Kamachi, M., & Gyoba, J. (1998, April). *Coding facial expressions with Gabor wavelets*. Paper presented at the Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan.

McAlonan, K., & Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioural Brain Research, 146*, 97–103.

McIntosh, A. R., Grady, C. L., Haxby, J. V., Ungerleider, L. G., & Horwitz, B. (1996). Changes in limbic and prefrontal functional interactions in a working memory task for faces. *Cerebral Cortex, 6*, 571–584.

Meunier, M., Bachevalier, J., & Mishkin, M. (1997). Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. *Neuropsychologia, 35*, 999–1015.

Newmark, R. E., Schon, K., Ross, R. S., & Stern, C. E. (2013). Contributions of the hippocampal subfields and entorhinal cortex to disambiguation during working memory. *Hippocampus*. doi:10.1002/hipo.22106

Nichols, E. A., Kao, Y. C., Verfaellie, M., & Gabrieli, J. D. E. (2006). Working memory and long-term memory for faces: Evidence from fMRI and global amnesia for involvement of the medial temporal lobes. *Hippocampus, 16*, 604–616.

Olsen, R. K., Nichols, E. A., Chen, J., Hunt, J. F., Glover, G. H., Gabrieli, J. D. E., & Wagner, A. D. (2009). Performance-related sustained and anticipatory activity in human medial temporal lobe during delayed match-to-sample. *Journal of Neuroscience*, *29*, 11880–11890.

Olson, I. R., Moore, K. S., Stark, M., & Chatterjee, A. (2006). Visual working memory is impaired when the medial temporal lobe is damaged. *Journal of Cognitive Neuroscience, 18*, 1087–1097.

Olson, I. R., Page, K., Moore, K. S., Chatterjee, A., & Verfaellie, M. (2006). Working memory for conjunctions relies on the medial temporal lobe. *Journal of Neuroscience, 26*, 4596–4601.

Ongur, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *The Journal of Comparative Neurology, 460*, 425–449.

Otto, T., & Eichenbaum, H. (1992). Complementary roles of the orbital prefrontal cortex and the perirhinal–entorhinal cortices in an odor-guided delayed-nonmatching-to-sample task. *Behavioral Neuroscience, 106*, 762–775.

Pantic, M., Valstar, M. F., Rademaker, R., & Maat, L. (2005, July). *Web-based database for facial expression analysis*. Paper presented at the IEEE Int'l Conference on Multimedia and Expo, Amsterdam, The Netherlands.

Petrides, M. (2005). Lateral prefrontal cortex: Architectonic and functional organization. *Philosophical Transactions of the Royal Society B, 360*, 781–795.

Ranganath, C., Cohen, M. X., & Brozinsky, C. J. (2005). Working memory maintenance contributes to long-term memory formation: Neural and behavioral evidence. *Journal of Cognitive Neuroscience, 17*, 994–1010.

Ranganath, C., & D'Esposito, M. (2001). Medial temporal lobe activity associated with active maintenance of novel information. *Neuron, 31*, 865–873.

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage, 23*, 752–763.

Roberts, A. C., Tomic, D. L., Parkinson, C. H., Roeling, T. A., Cutter, D. J., Robbins, T. W., & Everitt, B. J. (2007). Forebrain connectivity of the prefrontal cortex in the marmoset monkey (*Callithrix jacchus*): An anterograde and retrograde tract-tracing study. *Journal of Comparative Neurology*, *502*, 86–112.

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain and Cognition, 55*, 11–29.

Rolls, E. T. (2007). The representation of information about faces in the temporal and frontal lobes. *Neuropsychologia, 45*, 124–143.

Ross, R. S., Brown, T. I., & Stern, C. E. (2009). The retrieval of learned sequences engages the hippocampus: Evidence from fMRI. *Hippocampus, 19*, 790–799.

Ross, R. S., Sherrill, K. R., & Stern, C. E. (2011). The hippocampus is functionally connected to the striatum and orbitofrontal cortex during context dependent decision making. *Brain Research, 1423*, 53–66.

Rudebeck, P. H., & Murray, E. A. (2008). Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *Journal of Neuroscience, 28*, 8338–8343.

Sala, J. B., Rama, P., & Courtney, S. M. (2003). Functional topography of a distributed neural system for spatial and nonspatial information maintenance in working memory. *Neuropsychologia, 41*, 341–356.

Scheperjans, F., Eickhoff, S. B., Hömke, L., Mohlberg, H., Hermann, K., Amunts, K., & Zilles, K. (2008). Probabilistic maps, morphometry, and variability of cytoarchitectonic areas in the human superior parietal cortex. *Cerebral Cortex*, *18*, 2141–2157.

Schluppeck, D., Curtis, C. E., Glimcher, P. W., & Heeger, D. J. (2006). Sustained activity in topographic areas of human posterior parietal cortex during memory-guided saccades. *Journal of Neuroscience, 26*, 5098–5108.

Schoenbaum, G., Setlow, B., Nugent, S. L., Saddoris, M. P., & Gallagher, M. (2003). Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learning and Memory, 10*, 129–140.

Schon, K., Atri, A., Hasselmo, M. E., Tricarico, M. D., LoPresti, M. L., & Stern, C. E. (2005). Scopolamine reduces persistent activity related to long-term encoding in the parahippocampal gyrus during delayed matching in humans. *Journal of Neuroscience, 25*, 9112–9123.

Schon, K., Hasselmo, M. E., Lopresti, M. L., Tricarico, M. D., & Stern, C. E. (2004). Persistence of parahippocampal representation in the absence of stimulus input enhances long-term encoding: A functional magnetic resonance imaging study of subsequent memory after a delayed match-to-sample task. *Journal of Neuroscience, 24*, 11088–11097. doi:10.1523/JNEUROSCI.3807-04.2004

Schon, K., Quiroz, Y. T., Hasselmo, M. E., & Stern, C. E. (2009). Greater working memory load results in greater medial temporal activity at retrieval. *Cerebral Cortex, 19*, 2561–2571.

Schon, K., Ross, R. S., Hasselmo, M. E., & Stern, C. E. (2013). Complementary roles of medial temporal lobes and middorsolateral prefrontal cortex for working memory for novel and familiar trial-unique visual stimuli. *European Journal of Neuroscience, 37*, 668–678. doi:10.1111/ejn.12062

Schon, K., Tinaz, S., Somers, D. C., & Stern, C. E. (2008). Delayed match to object or place: An event-related fMRI study of short-term stimulus maintenance and the role of stimulus pre-exposure. *NeuroImage, 39*, 857–872.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal–midbrain encoding of overlapping events. *Neuron, 60*, 378–389.

Shrager, Y., Levy, D. A., Hopkins, R. O., & Squire, L. R. (2008). Working memory and the organization of brain systems. *Journal of Neuroscience, 28*, 4818–4822.

Slotnick, S. D., Moo, L. R., Segal, J. B., & Hart, J., Jr. (2003). Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Cognitive Brain Research, 17*, 75–82.

Stern, C. E., Sherman, S. J., Kirchhoff, B. A., & Hasselmo, M. E. (2001). Medial temporal and prefrontal contributions to working memory tasks with novel and familiar stimuli. *Hippocampus, 11*, 337–346.

Tabbert, K., Stark, R., Kirsch, P., & Vaitl, D. (2005). Hemodynamic responses of the amygdala, the orbitofrontal cortex and the visual cortex during a fear conditioning paradigm. *International Journal of Psychophysiology, 57*, 15–23.

Tsuchida, A., Doll, B. B., & Fellows, L. K. (2010). Beyond reversal: A critical role for human orbitofrontal cortex in flexible learning from probabilistic feedback. *Journal of Neuroscience, 30*, 16868–16875.

Walton, M. E., Behrens, T. E., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron, 65*, 927–939.

Wood, E. R., Dudchenko, P. A., Robitsek, R. J., & Eichenbaum, H. (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron, 27*, 623–633.