Research

Neural correlates of the encoding of multimodal contextual features

Lauren J. Gottlieb,[1,3] Jenny Wong,[2] Marianne de Chastelaine,[2] and Michael D. Rugg[2,4]

[1]Center for the Neurobiology of Learning and Memory, Department of Neurobiology and Behavior, University of California at Irvine, Irvine, California 92697, USA; [2]Center for Vital Longevity, School of Behavioral and Brain Sciences, The University of Texas at Dallas, Dallas, Texas 75235, USA

Functional magnetic resonance imaging (fMRI) was employed to identify neural regions engaged during the encoding of contextual features belonging to different modalities. Subjects studied objects that were presented to the left or right of fixation. Each object was paired with its name, spoken in either a male or a female voice. The test requirement was to discriminate studied from unstudied pictures and, for each picture judged old, to retrieve its study location and the gender of the voice that spoke its name. Study trials associated with accurate rather than inaccurate location memory demonstrated enhanced activity in the fusiform and parahippocampal cortex and the hippocampus and reduced activity (a negative subsequent memory effect) in the medial occipital cortex. Successful encoding of voice information was associated with enhanced study activity in the right middle superior temporal sulcus and activity reduction in the right superior frontal cortex. These findings support the proposal that encoding of a contextual feature is associated with enhanced activity in regions engaged during its online processing. In addition, they indicate that negative subsequent memory effects can also demonstrate feature-selectivity. Relative to other classes of study trials, trials for which both contextual features were later retrieved demonstrated enhanced activity in the lateral occipital complex and reduced activity in the temporo-parietal junction. These findings suggest that multifeatural encoding was facilitated when the study item was processed efficiently and study processing was not interrupted by redirection of attention toward extraneous events.

Episodic memories—memories for unique events—depend on the ability to associate or bind the various components of an event into a common memory representation. In laboratory studies, experimentally manipulated components of an event frequently include an "item" that is the focus of attention and one or more contextual features. Memory is tested using a source memory procedure in which, for each correctly recognized studied item, subjects are required to retrieve its associated contextual feature(s). Successful source judgments are assumed to be indicative of memory representations for which item and context information were successfully associated or bound together at the time of encoding.

A sizeable literature has developed in which event-related functional magnetic resonance imaging (fMRI) was employed to identify the neural correlates of successful encoding of item-context associations through the use of the "subsequent memory procedure" (Paller and Wagner 2002). In these studies, successful source encoding is consistently reported to be associated with enhanced activity (relative to study items for which later source retrieval failed) in the hippocampus and adjacent regions of the medial temporal lobe (MTL) (e.g., Davachi et al. 2003; Ranganath et al. 2004; Kensinger and Schacter 2006; Staresina and Davachi 2006, 2008; Uncapher et al. 2006; Uncapher and Rugg 2009; Diana et al. 2010; for reviews, see Davachi 2006; Diana et al. 2007; Kim 2011; see Gold et al. 2006 and Kirwan et al. 2008 for an alternative interpretation of the findings). The findings converge with evidence from neuropsychological and animal studies to suggest that the formation of item-context asso-

ciations in the course of a single study trial is especially dependent upon the MTL and, most notably, the hippocampus (Eichenbaum et al. 2007).

Less is known about the contribution of cortical regions to the encoding of item-context associations. In one prominent class of models (e.g., Alvarez and Squire 1994; Rolls 2000; Shastri 2002; Norman and O'Reilly 2003; for review, see Rugg et al. 2008), the hippocampus supports episodic memory encoding by capturing the patterns of cortical activity elicited during the online processing of an event and binding the patterns into a common memory representation. According to this framework, the cortical correlates of the encoding of a given event should overlap with the regions engaged during its online processing. The findings of several fMRI subsequent memory studies support this proposal (Otten and Rugg 2001; Otten et al. 2002; Fletcher et al. 2003; Mitchell et al. 2004; Uncapher et al. 2006; Otten 2007; Park and Rugg 2008; Park et al. 2008; Uncapher and Rugg 2009; Gottlieb et al. 2010; Kuhl et al. 2012). Below, we focus on three studies from our laboratory that investigated the neural correlates of encoding different classes of contextual features.

In the study of Uncapher et al. (2006), two orthogonally varying features (font color and location) were associated with each study item. Subsequent memory effects predicting later memory for item-color associations were localized to a different cortical region than the regions where activity predicted memory for item-location associations. The regions demonstrating these color- and location-selective encoding effects overlapped regions implicated by prior studies in the processing of color and location information, respectively. The findings are consistent with the prediction that cortical subsequent memory effects reflect modulation of activity supporting the online processing of a study episode (see above). Uncapher et al. (2006) proposed that feature-

selective subsequent memory effects are a consequence of fluctuations in the allocation of attentional resources to the feature. They suggested that as the allocation of resources to a given feature increases, so does activity in the cortical regions engaged by the feature. The enhanced activity strengthens the feature's online representation and increases the probability that the associated pattern of cortical activity will be bound into a hippocampally mediated episodic memory representation. This proposal received direct support from the study of Uncapher and Rugg (2009), in which it was reported that feature-selective subsequent memory effects were, indeed, enhanced when attention was directed toward the relevant feature.

Whereas the two studies described above employed visual contextual features (color and location), Gottlieb et al. (2010) contrasted the subsequent memory effects associated with visual and auditory contextual information. Subjects studied picture-name pairs, with the name of the picture presented either visually or auditorily. The test requirement was to discriminate between unstudied pictures, studied pictures paired with a visual word, and studied pictures paired with an auditory word. Subsequent memory effects predictive of later memory for the visual and auditory words overlapped regions selectively engaged by study trials containing visual versus auditory words, respectively. Thus, the principle that encoding of a contextual feature is associated with relative enhancement of processing in cortical regions engaged during the online processing of the feature extends to the auditory modality.

In addition to identifying the neural correlates of successful encoding of the contextual features of location and color, Uncapher et al. (2006) also investigated the neural correlates of successful multifeatural encoding, that is, the subsequent memory effects elicited by study trials for which both contextual features were successfully retrieved on the later memory test. They reported that, uniquely, these study trials elicited a subsequent memory effect in the right intra-parietal sulcus (IPS), in the vicinity of a region previously implicated in object-based attention (Treisman 1998). Uncapher et al. (2006) interpreted this finding as evidence that later memory for both features was enhanced when the encoded information took the form of an integrated perceptual representation, rather than distinct item and contextual features. There was, however, no sign of such an effect in the study of Uncapher and Rugg (2009). In that study, though, subjects were required to attend to only one of the two features on each trial, perhaps reducing the likelihood that features would be bound into a common perceptual representation. Another potentially relevant factor is that whereas in the original experiment color was an intrinsic feature of the study items (font color of visually presented words), the study items in Uncapher and Rugg (2009) were pictures, with the color of the surrounding border constituting the relevant contextual feature. It is therefore possible that the failure to find a unique multifeatural subsequent memory effect in that study was because one of the features was not intrinsic to the object and hence was not incorporated into an object-centered representation of the study item.

The present experiment takes as its starting point the studies of Uncapher et al. (2006) and Gottlieb et al. (2010). Each study item (a picture of an object) was associated with two orthogonally varying, temporally discontiguous contextual features that belonged to different modalities and processing domains (spatial location and speaker voice). The study had three aims. The first was to replicate and extend the findings of Gottlieb et al. (2010), testing the prediction that voice-selective subsequent memory effects would be evident in regions of lateral temporal cortex implicated in the processing of voice information, whereas location-selective effects would be identified in regions supporting the processing of spatial information. Note that the first of these predictions in particular is significantly more specific than the predictions motivating our prior study (Gottlieb et al. 2010). In the present case, subsequent memory effects are predicted not merely in the auditorily responsive cortex, but in regions—such as the middle superior temporal sulcus (STS)—that support voice identification (Kriegstein and Giraud 2004; Belin 2006; Blank et al. 2011). Second, we addressed the question of whether successful conjoint encoding of the two contextual features would be associated with cortical subsequent memory effects additional to those elicited by the encoding of single features (cf. Uncapher et al. 2006). Such a finding would constitute a significant extension of the results reported by Uncapher et al. (2006) and suggest a role for the cortex as well as the medial temporal lobe in the binding of visual and auditory contextual features into a common memory representation. The failure to find such an effect would, however, be consistent with the findings reported by Uncapher and Rugg (2009) and suggest that unique multifeatural subsequent memory effects are limited to circumstances where the different features can be conjoined into a common perceptual representation (however, see the Discussion for an alternative account of Uncapher and Rugg's findings).

The third aim of the present study was to investigate the relationship between the encoding of item-context associations and "negative" subsequent memory effects—effects that take the form of relatively lower activity for later-remembered than later-forgotten study items (for reviews, see Uncapher and Wagner 2009; Kim 2011). Although they are reported relatively infrequently, negative subsequent memory effects are highly robust and, in terms of their spatial extent, can overshadow the more commonly investigated "positive" effects discussed above (e.g., Park and Rugg 2008). They are often held to result from modulation of "default mode" activity (Gusnard and Raichle 2001; Buckner et al. 2008), reflecting the benefit to encoding that accrues when processing resources are fully disengaged from internally directed cognition and allocated to a study episode (Daselaar et al. 2004; Kim 2011). There have been few efforts to directly investigate the functional significance of negative subsequent memory effects, however (but see Uncapher et al. 2011), and to our knowledge, there have been no prior reports of negative effects for the encoding of item-context associations. Therefore, it remains to be established whether these effects are sensitive to the encoding of such associations and, if so, whether the effects vary according to the type of contextual feature retrieved. Importantly, if negative subsequent memory effects for different features are anatomically dissociable, it would suggest that the effects reflect more than the modulation of a single functional network. Thus, we took advantage of the design of the present study to address the question whether negative subsequent memory effects differ according to the type and number of contextual features that were successfully encoded.

Results

Behavioral performance

Study task

Mean accuracy of the congruency judgments (see Materials and Methods) on study trials was 0.98 (SD = 0.02). Study reaction times (RTs), timed from the onsets of the study words, are given in Table 1, segregated according to subsequent memory performance. To parallel the fMRI analyses described below, RT analyses were conducted on the four subsequent memory conditions, "both correct," "location only," "voice-only," and "miss." A one-way ANOVA revealed no significant differences between the conditions ($F_{(2.6, 41.2)} = 1.25$).

Table 1. Mean reaction times (msec) for congruency judgments at study segregated by subsequent memory at test (SD in parentheses)

Class of subsequent memory effect	Reaction time (msec)
Both (location and voice)	1000 (207)
Location only	1002 (214)
Voice only	974 (218)
Item only	964 (231)
Miss	1000 (251)

Test task

Collapsed over source accuracy, the item hit rate was 0.83 (SD = 0.12) against a false alarm rate of 0.03. Mean proportions of responses in the different subsequent memory conditions are shown in Table 2. Source recollection was estimated with an index derived from a single high-threshold model, in which the probability of recollection was computed as: $\{p$ (source hit) $- 0.5$ $[1 -$ (source unsure)$]\}/\{1 - 0.5$ $[1 - p$ (source unsure)$]\}$. The index was calculated separately for location and voice. For the location condition, "source hit" refers to studied pictures that were recognized and assigned to a correct location judgment regardless of voice accuracy, and "source unsure" refers to recognized pictures followed by an "unsure" location judgment. The analogous assignments apply to the voice condition. Source recollection estimates were 0.68 (SD = 0.10) and 0.40 (SD = 0.17) for the location and voice conditions, respectively. These values significantly differed from one another ($t_{(16)} = 6.42$, $P < 0.001$) and, in each case, from the chance value of zero (location: $t_{(16)} = 27.82$, $P < 0.001$; voice: $t_{(16)} = 9.66$, $P < 0.001$).

fMRI results

Whereas all subjects contributed sufficient study trials to the three source retrieval conditions (means [ranges] of 79 [40–108], 32 [14–62], and 13 [7–24][5] for both correct, location only, and voice-only conditions, respectively), the great majority of subjects had too few trials in one or more of the remaining trial types (item only and item miss) to allow stable estimates of the activity elicited by these different trial types. Therefore, as in Gottlieb et al. (2010), the two trial types were collapsed to form a single category of "miss" trials containing all study items for which source-specifying information was unavailable. Consequently, the analyses reported below identify the neural correlates of the encoding of single and multiple contextual features, but they do not speak to the question of the neural correlates of item recognition in the absence of source-specifying information. The mean (range) trial number for the miss condition was 35 (12–77).

We first identified subsequent memory effects that were shared across the three critical trial types (both correct, location only, and voice only). In three additional analyses, we then identified effects that were associated selectively with each trial type

(cf. Uncapher et al. 2006). For illustrative purposes (Kriegeskorte et al. 2009), parameter estimates for the different classes of study trial are shown for the effects preferentially associated with the encoding of location and voice information, as well as for trials associated with the conjoint encoding of both features.

Common subsequent memory effects

Common effects were identified by the weighted contrast of both, location only, voice-only, and miss study trials (weights of 1, 1, 1, and −3, respectively, thresholded at $P < 0.001$ with a 21-voxel cluster extent threshold), inclusively masked with the both > miss, location only > miss, and voice-only > miss contrasts (each thresholded at $P < 0.05$). Thus, this contrast identified voxels demonstrating both a main effect of subsequent memory and, additionally, a significant effect for each class of study trial. The outcome of this procedure is illustrated in Figure 1 and described in Table 3. Identified regions included posterior and middle aspects of the right superior temporal sulcus, right posterior hippocampus, left amygdala, and the caudate nucleus.

Feature–sensitive effects

Subsequent memory effects preferentially associated with later memory for location were identified by inclusively masking the contrast between location only and miss study trials (thresholded at $P < 0.001$) with the location only > voice-only contrast ($P < 0.05$), thereby identifying voxels where subsequent location effects exceeded subsequent voice effects. The analogous contrasts were employed to identify voxels demonstrating voice-selective effects. The outcome of each analysis is illustrated in Figure 2 and described in Table 3. Location effects were identified in the left ventral fusiform and parahippocampal cortex and the left anterior medial temporal lobe, including the hippocampus. Voice effects were confined to a single cluster located in the middle STS.

Conjoint effects

Subsequent memory effects sensitive to the conjoint retrieval of location and voice information were identified by masking the contrast between both and miss study trials ($P < 0.001$) with the both > location only and both > voice-only contrasts ($P < 0.05$). As is illustrated in Figure 3, this analysis revealed a single cluster in the right lateral occipital cortex (peak voxel: 42, −70, −5, $Z = 4.20$, 22 voxels).
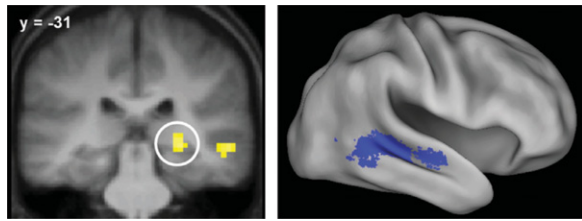
Unlike in Uncapher et al. (2006), the foregoing analysis failed to identify conjoint subsequent memory effects in the parietal cortex. In a final analysis, we performed a small volume correction for the both > miss contrast within a 5-mm-radius sphere centered on the peak voxel in the IPS where Uncapher et al. (2006) reported conjoint effects (21, −42, 48). No effects were identified.

---

[5]Data were accepted from subjects who contributed a minimum of seven trials to each critical experimental condition. As is evident, the voice-only condition was associated with fewer trials than the remaining conditions. These relatively low trial numbers raise the possibility that parameter estimates associated with this condition may have been less stable or reliable than those associated with the remaining conditions. This does not appear to have been the case, however. First, the across-subject variance of the parameter estimates from the four critical conditions did not systematically vary, suggesting that the stability of the parameter estimates was not substantially influenced by trial number. Second, a reanalysis of the fMRI data that omitted the five subjects with fewer than 10 trials in the voice-only condition revealed a pattern of effects that, while less robust statistically, was qualitatively almost identical to that reported in the Results section.

Table 2. Source memory performance conditionalized on items receiving a correct recognition judgment at retrieval

Study trial	Proportion correct
Both (location and voice)	0.60 (0.10)
Location only	0.24 (0.08)
Voice only	0.10 (0.04)
Item only	0.06 (0.04)

Mean proportions of correct responses are given for class of study trial (SD in parentheses).

Figure 1. Common subsequent memory effects in the right posterior hippocampus (*left*) and right posterior and middle superior temporal sulcus (*right*). Results are overlaid onto a coronal section of the across-subjects mean T1-weighted anatomical image and the standardized brain of the PALS-B12 atlas implemented in Caret5 (Van Essen 2005).

Negative subsequent memory effects

We employed analysis procedures analogous to those described above to identify common, feature-sensitive, and conjoint negative subsequent memory effects. These effects are illustrated in Figure 4 and listed in Table 4. Common effects were identified in medial and lateral parietal and bilateral orbital prefrontal cortex. Voice- and location-sensitive effects were identified in the right superior frontal sulcus and medial occipito-parietal cortex, respectively (the latter just posterior to the common-effect-identified medial parietal cortex). Finally, conjoint negative subsequent memory effects were identified in the left and right temporo-parietal junction (TPJ), anterior to the lateral parietal clusters demonstrating common effects.

Discussion

The aim of the present study was to shed light on how contextual features belonging to different sensory modalities are bound with item information to form an accessible episodic memory representation. Thus, we contrasted the neural correlates of the incidental encoding of associations between pictures of objects, their spatial location, and the gender in which the name of the picture was spoken. We identified four classes of subsequent memory effect: effects common to the encoding, either singly or conjointly, of both features, effects preferentially associated with the encoding of either location or voice information, and effects preferentially associated with the encoding of both features rather than a single feature. We also identified analogous negative subsequent memory effects. Below, we discuss the implications of these findings for the understanding of the encoding of item-context associations.

Behavioral findings

Study RTs did not significantly vary between the different subsequent memory conditions. Thus, it is unlikely that the different subsequent memory effects discussed below reflected gross differences in the efficiency with which the respective classes of study events were processed. Source memory was, however, higher for location than for voice information. Whereas the impact of this disparity should have been mitigated by the employment of a "source unsure" condition to minimize the diluting influence of guesses on subsequent memory effects (see also Park et al. 2008), it remains possible that some location-sensitive effects reflected differences in the strengths of the memories supporting location and voice information, rather than in the nature of what was encoded (cf. Squire et al. 2007; Kirwan et al. 2008). This caveat does not, however, apply to the interpretation of voice-sensitive subsequent memory effects. Together with prior evidence that feature-sensitive subsequent memory effects are dissociable from memory

strength (Uncapher et al. 2006; Uncapher and Rugg 2009), this finding strongly suggests that differential feature-sensitive effects reflect the encoding of qualitatively different kinds of information.

fMRI findings

Common effects

Subsequent memory effects common to study items receiving at least one correct source judgment were evident in the posterior hippocampus and amygdala, the STS, and the caudate nucleus. The hippocampal effects replicate numerous prior reports linking subsequent memory effects in this region to encoding processes that support later recollection of study details (see Introduction). The findings are consistent with the widely held view that the hippocampus contributes to the formation of associations between different components of an event ("memory binding") such as a study item and its context (e.g., Davachi 2006). Subsequent memory effects in the amygdala have frequently been reported to be more prominent for emotionally arousing study materials than for the kinds of materials employed in the present study (for reviews, see Kensinger 2004, 2009), findings consistent with the large body of evidence implicating the amygdala in the emotional modulation of memory (McGaugh 2003). A recent metaanalysis (Kim 2011) revealed, however, that amygdala subsequent memory effects are a consistent finding in studies employing nonemotional materials and that these effects are larger in studies testing "associative" memory (source memory or memory for item–item associations) than in studies testing memory for single items, at least in the case of pictorial study materials. The present results are consistent with these prior findings and, together with that body of research, raise the possibility that the amygdala plays a broader role in episodic memory than is often assumed.

Table 3. Loci of subsequent memory effects

Coordinates			Peak Z	Number of above-threshold voxels	Region
x	y	z			
Common effects					
−24	2	−26	3.65	24	L amygdala
−39	−97	1	3.40	22	L occipital cortex
9	2	19	3.65	81	R head of caudate
18	23	10	4.41	43	R head of caudate
54	−7	−8	3.72	30	R superior temporal sulcus
33	−43	4	4.12	232	R hippocampal tail
Subpeaks					
27	−31	−2	3.68		R posterior hippocampus
51	−43	1	3.78		R posterior superior temporal sulcus
Location-only effects					
−48	−58	−14	4.20	246	L fusiform
Subpeaks					
−21	−10	−23	3.85		L anterior hippocampus
−30	−43	−26	3.82		L parahippocampal cortex
Voice-only effects					
54	−7	−8	4.25	54	R middle superior temporal sulcus
Conjoint effects					
42	−70	−5	4.20	22	R lateral occipital cortex
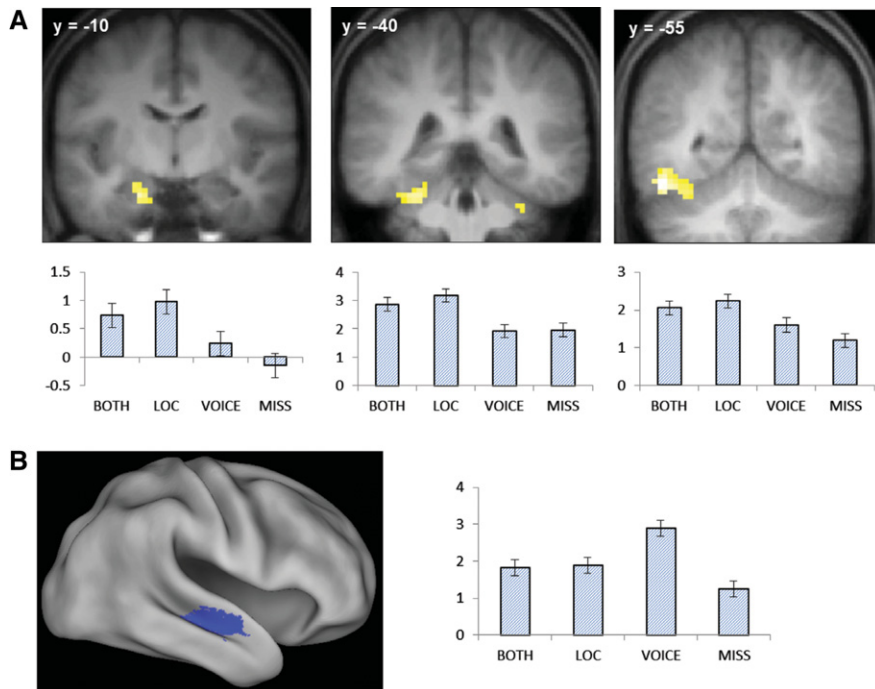
Figure 2. (*A*) Location-sensitive subsequent memory effects. Bar plots show (*left* to *right*) parameter estimates for the different classes of study trial for voxels in the anterior hippocampus (-21, -10, -23), parahippocampal cortex (-30, -40, -26), and fusiform gyrus (-48, -55, -14). Error bars here and in the following figures signify the standard error of the mean derived from the error term of the one-way ANOVA (see text and Loftus and Masson 1994). (*B*) Voice-sensitive subsequent memory effect. Bar plot shows peak parameter estimates for the different classes of study trial (54, -7, -8).

Subsequent memory effects common to the encoding of location and voice information were also evident in the right STS, primarily in its posterior aspect, which has been implicated in the processing of complex verbal and nonverbal sounds (Kriegstein and Giraud 2004). In light of the findings from our prior study (Gottlieb et al. 2010), in which auditory subsequent memory effects were reported in an STS region overlapping that identified here (peak voxels of 60, -9, -15 and 54, -7, -8, respectively), it was expected that the successful encoding of voice gender would be associated with enhanced study activity in the STS (see below). It is less obvious, though, why items for which such encoding failed (location only items) also elicited STS subsequent memory effects. A possible explanation lies in the specificity of the voice information that had to be encoded to support accurate performance on the later source memory test. The information had to support memory not merely for whether a name had been heard (as in our prior study) but for the gender of the voice in which the name was spoken. We conjecture that on some proportion of the trials where location information was successfully encoded, encoding also encompassed auditory features of the associated word that were nondiagnostic of gender. Under this scenario, location-only subsequent memory effects would be accompanied by effects reflecting the encoding of this nondiagnostic auditory information.

As with the amygdala, the caudate nucleus also is not usually considered to play a role in episodic memory. However, subsequent memory effects in the caudate are not without precedent (Ben-Yakov and Dudai 2011; Blumenfeld et al. 2011), although not reported with sufficient consistency to be identified in the meta-analysis of Kim (2011). Along with effects in the dorsolateral prefrontal cortex, caudate subsequent memory effects have been

reported to be especially prominent for the encoding of temporally discontiguous associations (Qin et al. 2007). The present findings can, perhaps, be understood in this light, since visual and auditory information were presented consecutively rather than concurrently. Even at a reduced threshold, however, subsequent memory effects could not be identified in dorsolateral prefrontal cortex.

Feature–sensitive subsequent memory effects

Location-only subsequent memory effects were enhanced relative to voice-only effects in left ventral fusiform and parahippocampal cortex and throughout the left MTL, including the hippocampus. The finding for the left fusiform and parahippocampal cortex replicates previous reports of subsequent source memory effects in these regions for the encoding of location (Cansino et al. 2002; Sommer et al. 2005a,b; Park et al. 2008; Uncapher and Rugg 2009). As already noted, some caution in interpreting these effects is warranted, given the disparity in source memory accuracy for location and voice. Nonetheless, the finding that successful location encoding was associated with enhanced parahippocampal activity is consistent with proposals that this region plays a key role in representing spatial context and transmitting this information to the hippocampus (Davachi 2006; Diana et al. 2007; Eichenbaum et al. 2011). In contrast, successful encoding of voice information was associated with an enhanced subsequent memory effect in the middle STS. This region has previously been implicated in the processing of high-level vocal features such as speaker identity and gender (Kriegstein and Giraud 2004; Belin 2006; Blank et al. 2011).
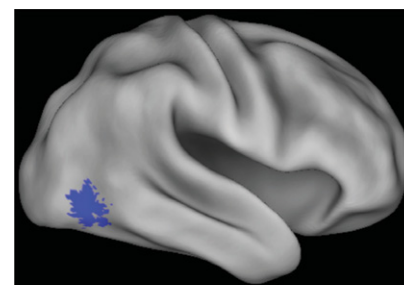
Figure 3. Subsequent memory effect for the retrieval of both location and voice. Bar plot shows peak parameter estimates for the different classes of study trial (42, -70, -5).
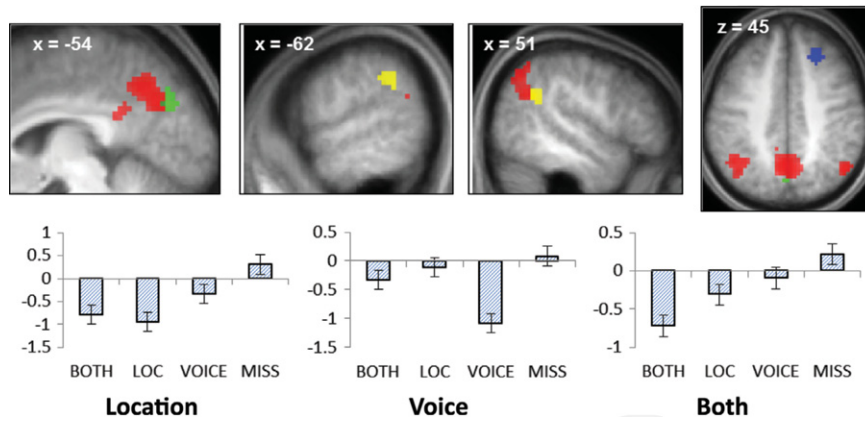
Figure 4. Common (red), location-sensitive (green), voice-sensitive (blue), and both (yellow) negative subsequent memory effects. Bar plots show peak parameter estimates for each of the feature-sensitive effects (averaged across the left and right hemispheres in the case of the study trials for which both features were later retrieved).

The present findings of distinct location- and voice-sensitive feature-sensitive subsequent memory effects offer further support for the proposal that subsequent memory effects reflect enhancement of neural activity engaged during the online processing of a study episode (Rugg et al. 2008). As was discussed in the Introduction, this enhancement may depend on how attentional resources are allocated as an episode unfolds. By this account, memory for location benefited when subjects paid particular attention to the spatial layout of the study event, while memory for voice gender was facilitated when the processing of voice identity was emphasized. The present findings are consistent with this account but do not provide additional evidence in its support. This will require further studies along the lines of Uncapher and Rugg (2009), in which attention to different contextual features is directly manipulated.

The foregoing discussion leads to the question of whether any mechanism additional to that supporting the encoding of the individual contextual features is necessary to explain how the features were encoded conjointly. As was reviewed in the Introduction, Uncapher et al. (2006) reported that the conjoint encoding of the location and color of visually presented words, but not the encoding of either feature alone, was associated with subsequent memory effects in the right IPS. In light of the putative role of the IPS in object-centered attention (Treisman 1998), this finding was interpreted as evidence that the conjoint encoding of contextual features is facilitated when they are combined with item information to form an online representation in which the item and the features are integrated. Although caution is necessary when drawing a conclusion from a null result, the absence of an IPS subsequent memory effect in the present study is consistent with this proposal if it is assumed that temporally discontiguous visual and auditory information cannot be integrated into a common perceptual representation. Alternatively, the present null result might be a reflection of the role of the IPS in mediating attentional shifts between different features of a stimulus event (Schultz and Lennert 2009). By this argument,[6] the conjoint subsequent memory effects in the IPS reported by Uncapher et al. (2006) reflected a role for this region not in perceptual integration but in switching attention rapidly between the different contextual features, facilitating their conjoint encoding.

[6]We thank a reviewer of a previous version of this paper for bringing this possibility to our attention.

Such a mechanism would have been less relevant for conjoint encoding in the present case, however, since the spatial and auditory features were presented consecutively rather than concurrently.

Although conjoint subsequent memory effects could not be identified in the IPS, a conjoint effect was identified in the right lateral occipital cortex. This region, often referred to as the "lateral occipital complex" (LOC), is strongly implicated in the processing of object structure (Malach et al. 1995; Grill-Spector et al. 1999; Kourtzi and Kanwisher 2001; Reinholz and Pollmann 2005). The region was reported to demonstrate subsequent source memory effects (for picture-location associations) by Cansino et al. (2002). Subsequent memory effects in the LOC for the encoding of item-context and item-item associations have consistently been reported in subsequent studies that employed pictorial study materials (depictions of objects, scenes, or faces) (Kim 2011). The center of mass of these effects was estimated (Kim 2011) to be 42, −74, −4, close to the present peak of 42, −70, −5. It is unclear why enhanced LOC activity should be associated with more effective associative encoding of pictorial study items. One possibility is that the effects reflect the allocation of a relatively large amount of attentional resources to an item, facilitating its perceptual processing. From this perspective, therefore, the present findings suggest that one factor determining whether location and voice information were conjointly associated with a studied picture may have been the quality or richness of the object representation derived from the picture. A second possible factor is discussed below.

Negative subsequent memory effects

Robust negative subsequent memory effects were identified for the combination of all three trial types involving successful

Table 4. Loci of negative subsequent memory effects

Coordinates			Peak Z	Number of above-threshold voxels	Region
x	y	z			
Common effects					
−30	56	−14	3.82	28	L orbitofrontal cortex
−3	−58	46	4.58	388	Medial parietal cortex
−42	−61	40	5.17	373	L posterior parietal cortex
39	56	−17	4.08	21	R orbitofrontal cortex
48	−64	46	4.37	110	R posterior parietal cortex
Location-sensitive effects					
−6	−73	34	4.09	29	Medial parietal/ occipital cortex
Voice-sensitive effects					
24	32	46	4.27	48	R superior frontal sulcus
Conjoint effects					
−63	−46	34	4.12	27	L temporo-parietal junction
51	−52	31	4.45	37	R temporo-parietal junction

contextual retrieval, along with effects preferentially associated with memory either for location, voice, or the two features conjointly. To our knowledge, this report is the first to describe negative subsequent source memory effects and to contrast these effects according to the nature of the retrieved contextual feature.

Common negative subsequent memory effects were evident in the medial and bilateral posterior parietal cortex, regions that have consistently been reported to demonstrate negative effects across a wide variety of stimulus materials and study and test procedures (Kim 2011). These regions are components of the default mode network. As was noted in the Introduction, negative subsequent memory effects in default mode regions are thought to reflect the benefit to encoding that comes with full rather than partial disengagement of internally directed cognition in response to a study trial (Daselaar et al. 2004; Kim 2011).

Common negative subsequent memory effects were accompanied by both location- and voice-sensitive effects. Location-sensitive effects were identified in a small medial parietal/occipital cluster just posterior to the medial parietal region demonstrating common effects. Voice-sensitive negative subsequent memory effects, by contrast, were identified in the right superior frontal sulcus, distant from any region demonstrating common effects (see Fig. 4). The functional significance of these findings is obscure. They do suggest, however, that negative subsequent memory effects reflect more than the modulation of generic processes—such as those supported by the "default-mode network"—that impact episodic encoding in a nonselective fashion. It will be of interest to ascertain whether the present dissociation between location- and voice-sensitive negative subsequent memory effects extends to other contextual features.

Relative to all of the other classes of study trial, successful conjoint encoding of location and voice information was associated with reduced activity in the TPJ. This region is held to be a key component of a "ventral attentional network" that supports bottom-up attention—the capture of attention by events outside the current attentional focus (Corbetta et al. 2008). It has further been proposed that sustained deactivation of the TPJ reflects the engagement of a "filtering" operation that prevents attention from being captured by task-irrelevant events (Shulman et al. 2007). In light of these proposals, Uncapher and Wagner (2009) suggested that negative subsequent memory effects in the TPJ reflect the benefit to encoding that results when attention remains focused on the study event and is not redirected toward other aspects of the environment during the course of a study trial (see Uncapher et al. 2011 for experimental findings consistent with this proposal). From this perspective, the present finding suggests that conjoint encoding was facilitated when attention remained fully focused on salient aspects of the study event. We conjecture that this sustained focus was responsible for the enhanced processing of study pictures reflected in the conjoint lateral occipital subsequent memory effect discussed above. Additionally, it may have provided the attentional resources necessary to allow both contextual features, rather than one only, to be bound into the memory representation of the study event.

Concluding comments

The present findings provide further support for the proposal that successful contextual encoding is associated with enhanced activity in cortical regions engaged during the online processing of contextual features. The findings further suggest that conjoint encoding of different contextual features is facilitated when attention remains strongly focused throughout a study trial and promotes especially efficient processing of the study item.

Materials and Methods

Subjects

Twenty-nine subjects consented to participate in the study. All reported themselves to be in good general health, right-handed, to have no history of neurological disease or contraindications for MR imaging, and to have attained fluency in English by age five. They were recruited from the University of California, Irvine (UCI) community and remunerated for their participation in accordance with the human subjects procedures approved by the Institutional Review Board of UCI. The data from 11 subjects were excluded from all analyses: Nine subjects[7] failed to demonstrate above chance subsequent source memory for at least one of the features (source recollection [Psr] < 0.1 against a chance score of 0, as estimated using an index derived from a single high-threshold model; see below). Two additional subjects were excluded for having too few "voice-only" trials (N's of 4 and 3, respectively). Data are reported from the remaining 18 subjects (12 female), who ranged in age from 18 to 26 yr.

Stimulus materials

Three hundred colored pictures of common objects were employed as the experimental materials (http://www.hemera.com/index.html). The names of the pictures had a mean written frequency between 1 and 100 counts per million (Kucera and Francis 1967). The spoken name of each picture was recorded in two voices, one male and the other female. The auditory files (mean duration = 650 msec, maximum = 1000 sec) were edited to a constant peak sound pressure and filtered to remove ambient noise (http://audacity.sourceforge.net).

Of the 300 pictures, 10 served as buffers (two at the beginning and end of each half of the study list, and two at the beginning of the test list), and 30 were used in the practice phases preceding the study and test sessions (see below). For each subject, 160 pictures and their corresponding names were randomly selected from the remainder of the pool to serve as study items. For 80 of these pictures, the associated name was spoken in the male voice, and for the remaining 80, the name was spoken in the female voice. Pictures paired with the male or female voice were each equally likely to appear to the left or right of fixation.

Twenty additional pictures were paired with the names of an alternate object so that the picture and name were incongruent. These incongruent stimuli were equally distributed between male and female voices, and left- and right-sided locations. The 160 congruent study stimuli were randomly intermixed with these 20 incongruent stimuli to create a 180-item study list (excluding buffer trials; see above). The remaining 80 pictures were presented as foils in the test phase.

At study, the pictures were back-projected onto a screen and viewed via a mirror mounted on the scanner headcoil. The pictures were displaced 4° either to the left or right of fixation and presented within a continuously displayed solid gray frame which subtended a visual angle of 5.4° × 5.4°. Auditory words were presented binaurally via MR compatible headphones. The volume was adjusted in the scanner to a comfortable listening level for each subject prior to scanning.

Test items were presented outside of the scanner on a computer monitor. The test lists comprised the 160 pictures paired with a congruent name at study, randomly intermixed with 80 unstudied pictures. The pictures and associated cues were displayed in central vision within a continuously present gray background that subtended 6.8° × 6.8°.

Experimental tasks and procedures

The experiment comprised a single study-test cycle.

---

[7]Six of these subjects were excluded because of poor memory for speaker voice, one subject was excluded because of poor location memory, and two were excluded because of poor memory for both features.

Study procedure

Instructions and practice were administered outside the scanner. The study phase proper consisted of the presentation of the 188-item study list. This was presented in two blocks separated by a brief rest period (approx. 1 min). Each block contained 80 congruent and 10 incongruent study trials, in each case divided equally between left and right locations and male and female voices. Each study trial began with the presentation of a red fixation character in the center of the display frame for 500 sec. The character was replaced by a picture displayed for 500 msec to either the left or the right side of fixation. Immediately after picture offset, a word was presented in either a male or a female voice, and a centrally presented black fixation character was presented for 2500 msec, completing the trial. Subjects were informed they would receive no warning as to the side of picture presentation or the gender of the voice speaking the word. The task requirement was to respond with one or other index finger depending whether the picture and the spoken name were congruent or incongruent. Hand of response was counterbalanced across subjects. The instructions placed equal emphasis on speed and accuracy. In each study block, the 4500-msec stimulus-onset-asynchrony (SOA) was randomly modulated by the addition of 40 null trials (Josephs and Henson 1999). The different types of study item were presented in pseudorandom order, with no more than three trials of one type (left male, right male, left female, right female, or null) occurring consecutively.

Test procedure

Following the completion of the second study block, subjects were removed from the scanner and taken to an adjacent testing room. Only then were they informed of the source memory test and administered instructions and a short practice test. Approximately 10 min elapsed between the completion of the second study block and the beginning of the memory test. Each test trial began with a red fixation character presented in the center of a gray frame for 500 sec, followed by the presentation for 500 msec of a centrally presented picture within a solid gray frame.

The test items consisted of the 160 critical study items (congruent trials) and 80 randomly interspersed unstudied (new) pictures; no more than three items of one type were presented consecutively. Instructions were to judge whether each picture was studied or unstudied and to indicate the decision with the left (old) or right (new) index finger. Subjects were instructed to indicate "new" when uncertain about an item's study status. If a picture was judged "new," the test advanced to the next trial (with a 1-inter-trial interval, during which a black fixation character was presented). If a picture was judged "old," two additional decisions were required. First, the prompt "Left, Right, Unsure?" appeared in black letters, signaling the need to indicate the picture's study location. The judgment was signaled by one of three button presses: left index finger for "Left," right index finger for "Right," and right middle finger when the location could not be retrieved and the judgment would merely be a guess (henceforth, an "Unsure" response). A second prompt then appeared which read "Male, Female, Unsure?" In response to this prompt, the requirement was to indicate the gender of the voice speaking the associated word. "Male" judgments were signaled with the left index finger, "Female" judgments with the right index finger, and "Unsure" judgments with the right middle finger. The test was self-paced and presented as a single block, taking approximately 20 min.

fMRI data acquisition

A Philips Achieva 3T MR scanner (Philips Medical Systems) was used to acquire T1-weighted anatomical images (240 × 240 matrix, 1-mm isotropic voxels, 160 slices, sagittal acquisition, 3D MP-RAGE sequence) and T2*-weighted echoplanar images (EPI) [80 × 79 matrix, 3-mm × 3-m in-plane resolution, axial acquisition, flip angle 70°, echo time (TE) 30 msec] optimized for blood oxygenation level-dependent (BOLD) contrast. The data were acquired using a sensitivity encoding (SENSE) reduction factor of 1.5

on an eight-channel parallel imaging headcoil. Each EPI volume comprised thirty 3-mm-thick axial slices oriented parallel to the AC-PC plane and separated by 1-mm gaps, positioned to give full coverage of the cerebrum and most of the cerebellum. Data were acquired in two sessions of 265 volumes each, with a repetition time (TR) of 2 sec/volume. Volumes within sessions were acquired continuously in an ascending sequential order. The 4.5-sec SOA allowed for an effective sampling rate of the hemodynamic response of 2z. The first five volumes of each session were discarded to allow equilibration of tissue magnetization.

fMRI data analysis

The data were analyzed using Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology) (Friston et al. 1994) implemented under Matlab2008b (The Mathworks Inc.). Functional images were subjected to a two-pass spatial realignment. Images were realigned to the first image, generating a mean image across sessions. In the second pass, the raw images were realigned to the generated mean image. The images were then subjected to reorientation, spatial normalization to a standard EPI template (based on the Montreal Neurological Institute [MNI] reference brain) (Cocosco et al. 1997) and smoothing with an 8-mm FWHM Gaussian kernel. Functional time series were concatenated across sessions.

Statistical analyses were performed on the study phase data in two stages of a mixed effects model. In the first stage, neural activity elicited by the study pictures was modeled by δ functions (impulse events) that coincided with the onset of each picture. The ensuing BOLD response was modeled by convolving the neural functions with a canonical hemodynamic response function (HRF) and its temporal and dispersion derivatives (Friston et al. 1998) to yield regressors in a general linear model (GLM) that modeled the BOLD response to each event-type.

For the reasons discussed in the Results section, the principal analyses were confined to four events of interest: recognized pictures for which both the location and voice were later remembered (both correct), recognized pictures associated with correct location judgments but incorrect voice judgments (location only), recognized pictures associated with correct voice judgments but incorrect location judgments (voice-only), and pictures associated with incorrect location and voice judgments or pictures that were incorrectly judged new (miss). A fifth category of trials comprised events of no interest and included buffer trials, incongruent stimuli, and trials associated with incorrect or omitted congruency judgments. Six regressors modeled movement-related variance (three rigid-body translations and three rotations determined from the realignment stage). Session-specific constant terms modeling the mean over scans in each session were also entered into the design matrix.

For each voxel, the functional time series was highpass-filtered to 1/128z and scaled within-session to yield a grand mean of 100 across voxels and scans. Parameter estimates for events of interest were estimated using a general linear model. Nonsphericity of the error covariance was accommodated by an AR(1) model, in which the temporal autocorrelation was estimated by pooling over supra-threshold voxels (Friston et al. 2002). The parameters for each covariate and the hyperparameters governing the error covariance were estimated using restricted maximum likelihood (ReML). Parameter estimates for the four conditions of interest (both correct, location only, voice-only, and miss conditions, respectively) were derived for each participant and carried forward to a second level group-wise analysis. In this analysis, individual participants' parameter estimates for the four conditions of interest were entered into a repeated-measures one-way ANOVA model, as implemented in SPM8. Planned contrasts assessing the different effects of interest were performed using the common error term derived from the ANOVA. Protection against Type I error was effected by using the "analysis of functional neuroimages" (AFNI) AlphaSim tool (http://afni.nimh.nih.gov/afni/AFNI_Help/AlphaSim.html) to estimate the minimum cluster size necessary for a whole-brain cluster-wise corrected significance level of $P < 0.05$ at a height-threshold of

$P < 0.001$. The critical value was 21 contiguous voxels. As described in the Results section, each planned contrast was inclusively masked with one or more additional contrasts to identify effects preferentially associated with a particular class of study trial. The 21-voxel extent threshold was maintained after application of the mask. The peak voxels of clusters exhibiting reliable effects are reported in MNI coordinates.

Acknowledgments

References

Alvarez P, Squire LR. 1994. Memory consolidation and the medial temporal lobe: A simple network model. *Proc Natl Acad Sci* **91:** 7041–7045.

Belin P. 2006. Voice processing in human and non-human primates. *Philos Trans R Soc Lond B Biol Sci* **361:** 2091–2107.

Ben-Yakov A, Dudai Y. 2011. Constructing realistic engrams: Poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. *J Neurosci* **31:** 9032–9042.

Blank H, Anwander A, von Kriegstein K. 2011. Direct structural connections between voice- and face-recognition areas. *J Neurosci* **31:** 12906–12915.

Blumenfeld RS, Parks CM, Yonelinas AP, Ranganath C. 2011. Putting the pieces together: The role of dorsolateral prefrontal cortex in relational memory encoding. *J Cogn Neurosci* **23:** 257–265.

Buckner RL, Andrews-Hanna JR, Schacter DL. 2008. The brain's default network: Anatomy, function, and relevance to disease. *Ann N Y Acad Sci* **1124:** 1–38.

Cansino S, Maquet P, Dolan RJ, Rugg MD. 2002. Brain activity underlying encoding and retrieval of source memory. *Cereb Cortex* **12:** 1048–1056.

Cocosco C, Kollokian V, Kwan RKS, Pike G, Evans A. 1997. BrainWeb: Online interface to a 3D MRI simulated brain database. *Neuroimage* **5:** S425. http://www.bic.mni.mcgill.ca/users/crisco/HBM97_abs/HBM97_abs.pdf.

Corbetta M, Patel G, Shulman GL. 2008. The reorienting system of the human brain: From environment to theory of mind. *Neuron* **58:** 306–324.

Daselaar SM, Prince SE, Cabeza R. 2004. When less means more: Deactivations during encoding that predict subsequent memory. *Neuroimage* **23:** 921–927.

Davachi L. 2006. Item, context and relational episodic encoding in humans. *Curr Opin Neurobiol* **16:** 693–700.

Davachi L, Mitchell JP, Wagner AD. 2003. Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci* **100:** 2157–2162.

Diana RA, Yonelinas AP, Ranganath C. 2007. Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends Cogn Sci* **11:** 379–386.

Diana RA, Yonelinas AP, Ranganath C. 2010. Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *J Cogn Neurosci* **22:** 1808–1818.

Eichenbaum H, Yonelinas AP, Ranganath C. 2007. The medial temporal lobe and recognition memory. *Annu Rev Neurosci* **30:** 123–152.

Eichenbaum H, Sauvage M, Fortin N, Komorowski R, Lipton P. 2011. Towards a functional organization of episodic memory in the medial temporal lobe. *Neurosci Biobehav Rev* **36:** 1597–1608.

Fletcher PC, Stephenson CME, Carpenter TA, Donovan T, Bullmorel ET. 2003. Regional brain activations predicting subsequent memory success: An event-related fMRI study of the influence of encoding tasks. *Cortex* **39:** 1009–1026.

Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ. 1994. Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Mapp* **2:** 189–210.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. 1998. Event-related fMRI: Characterizing differential responses. *Neuroimage* **7:** 30–40.

Friston KJ, Penny W, Phillips C, Kiebel S, Hinton G, Ashburner J. 2002. Classical and Bayesian inference in neuroimaging: Theory. *Neuroimage* **16:** 465–483.

Gold JJ, Smith CN, Bayley PJ, Shrager Y, Brewer JB, Stark CEL, Hopkins RO, Squire LR. 2006. Item memory, source memory, and the medial temporal lobe: Concordant findings from fMRI and memory-impaired patients. *Proc Natl Acad Sci* **103:** 9351–9356.

Gottlieb LJ, Uncapher MR, Rugg MD. 2010. Dissociation of the neural correlates of visual and auditory contextual encoding. *Neuropsychologia* **48:** 137–144.

Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzchak Y, Malach R. 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* **24:** 187–203.

Gusnard DA, Raichle ME. 2001. Searching for a baseline: Functional imaging and the resting human brain. *Nat Rev Neurosci* **2:** 685–694.

Josephs O, Henson RN. 1999. Event-related functional magnetic resonance imaging: Modelling, inference and optimization. *Philos Trans R Soc Lond B Biol Sci* **354:** 1215–1228.

Kensinger EA. 2004. Remembering emotional experiences: The contribution of valence and arousal. *Rev Neurosci* **15:** 241–251.

Kensinger EA. 2009. Remembering the details: Effects of emotion. *Emot Rev* **1:** 99–113.

Kensinger EA, Schacter DL. 2006. Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *J Neurosci* **26:** 2564–2570.

Kim H. 2011. Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *Neuroimage* **54:** 2446–2461.

Kirwan CB, Wixted JT, Squire LR. 2008. Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. *J Neurosci* **28:** 10541–10548.

Kourtzi Z, Kanwisher N. 2001. Representation of perceived object shape by the human lateral occipital complex. *Science* **293:** 1506–1509.

Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. 2009. Circular analysis in systems neuroscience: The dangers of double dipping. *Nat Neurosci* **12:** 535–540.

Kriegstein KV, Giraud A-L. 2004. Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage* **22:** 948–955.

Kucera H, Francis WN. 1967. *Computational analysis of present-day American English*. Brown University Press, Providence, RI.

Kuhl BA, Rissman J, Wagner AD. 2012. Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia* **50:** 458–469.

Loftus GR, Masson MEJ. 1994. Using confidence intervals in within-subject designs. *Psychon Bull Rev* **1:** 476–490.

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci* **92:** 8135–8139.

McGaugh JL. 2003. *Memory and emotion: The making of lasting memories*. Columbia University Press, New York.

Mitchell KJ, Johnson MK, Raye CL, Greene EJ. 2004. Prefrontal cortex activity associated with source monitoring in a working memory task. *J Cogn Neurosci* **16:** 921–934.

Norman KA, O'Reilly RC. 2003. Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychol Rev* **110:** 611–646.

Otten LJ. 2007. Fragments of a larger whole: Retrieval cues constrain observed neural correlates of memory encoding. *Cereb Cortex* **17:** 2030–2038.

Otten LJ, Rugg MD. 2001. Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cereb Cortex* **11:** 1150–1160.

Otten LJ, Henson RNA, Rugg MD. 2002. State-related and item-related neural correlates of successful memory encoding. *Nat Neurosci* **5:** 1339–1344.

Paller KA, Wagner AD. 2002. Observing the transformation of experience into memory. *Trends Cogn Sci* **6:** 93–102.

Park H, Rugg MD. 2008. Neural correlates of successful encoding of semantically and phonologically mediated inter-item associations. *Neuroimage* **43:** 165–172.

Park H, Uncapher MR, Rugg MD. 2008. Effects of study task on the neural correlates of source encoding. *Learn Mem* **15:** 417–425.

Qin S, Piekema C, Petersson KM, Han B, Luo J, Fernández G. 2007. Probing the transformation of discontinuous associations into episodic memory: An event-related fMRI study. *Neuroimage* **38:** 212–222.

Ranganath C, Yonelinas AP, Cohen MX, Dy CJ, Tom SM, D'Esposito M. 2004. Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia* **42:** 2–13.

Reinholz J, Pollmann S. 2005. Differential activation of object-selective visual areas by passive viewing of pictures and words. *Brain Res Cogn Brain Res* **24:** 702–714.

Rolls ET. 2000. Memory systems in the brain. *Annu Rev Psychol* **51:** 599–630.

Rugg MD, Johnson JD, Park H, Uncapher MR. 2008. Encoding-retrieval overlap in human episodic memory: A functional neuroimaging perspective. *Prog Brain Res* **169:** 339–352.

Schultz J, Lennert T. 2009. BOLD signal in intraparietal sulcus covaries with magnitude of implicitly driven attention shifts. *Neuroimage* **45:** 1314–1328.

Shastri L. 2002. Episodic memory and cortico-hippocampal interactions. *Trends Cogn Sci* **6:** 162–168.

Shulman GL, Astafiev SV, McAvoy MP, d' Avossa G, Corbetta M. 2007. Right TPJ deactivation during visual search: Functional significance and support for a filter hypothesis. *Cereb Cortex* **17:** 2625–2633.

Sommer T, Rose M, Gläscher J, Wolbers T, Büchel C. 2005a. Dissociable contributions within the medial temporal lobe to encoding of object-location associations. *Learn Mem* **12:** 343–351.

Sommer T, Rose M, Weiller C, Büchel C. 2005b. Contributions of occipital, parietal and parahippocampal cortex to encoding of object-location associations. *Neuropsychologia* **43:** 732–743.

Squire LR, Wixted JT, Clark RE. 2007. Recognition memory and the medial temporal lobe: A new perspective. *Nat Rev Neurosci* **8:** 872–883.

Staresina BP, Davachi L. 2006. Differential encoding mechanisms for subsequent associative recognition and free recall. *J Neurosci* **26:** 9162–9172.

Staresina BP, Davachi L. 2008. Selective and shared contributions of the hippocampus and perirhinal cortex to episodic item and associative encoding. *J Cogn Neurosci* **20:** 1478–1489.

Treisman A. 1998. Feature binding, attention and object perception. *Philos Trans R Soc Lond B Biol Sci* **353:** 1295–1306.

Uncapher MR, Rugg MD. 2009. Selecting for memory? The influence of selective attention on the mnemonic binding of contextual information. *J Neurosci* **29:** 8270–8279.

Uncapher MR, Wagner AD. 2009. Posterior parietal cortex and episodic encoding: Insights from fMRI subsequent memory effects and dual-attention theory. *Neurobiol Learn Mem* **91:** 139–154.

Uncapher MR, Otten LJ, Rugg MD. 2006. Episodic encoding is more than the sum of its parts: An fMRI investigation of multifeatural contextual encoding. *Neuron* **52:** 547–556.

Uncapher MR, Hutchinson JB, Wagner AD. 2011. Dissociable effects of top-down and bottom-up attention during episodic encoding. *J Neurosci* **31:** 12613–12628.

Van Essen DC. 2005. A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *Neuroimage* **28:** 635–662.