Voluntary Explicit versus Involuntary Conceptual Memory Are Associated with Dissociable fMRI Responses in Hippocampus, Amygdala, and Parietal Cortex for Emotional and Neutral Word Pairs

Cristina Ramponi[1], Philip J. Barnard[1], Ferath Kherif[2,3], and Richard N. Henson[1]

Abstract

■ Although functional neuroimaging studies have supported the distinction between explicit and implicit forms of memory, few have matched explicit and implicit tests closely, and most of these tested perceptual rather than conceptual implicit memory. We compared event-related fMRI responses during an intentional test, in which a group of participants used a cue word to recall its associate from a prior study phase, with those in an incidental test, in which a different group of participants used the same cue to produce the first associate that came to mind. Both semantic relative to phonemic processing at study, and emotional relative to neutral word pairs, increased target completions in the intentional test, but not in the incidental test, suggesting that behavioral performance in the incidental test was not contaminated by voluntary explicit retrieval. We isolated the neural correlates of successful retrieval by contrasting fMRI responses to studied versus unstudied cues for which the equivalent "target" associate was produced. By comparing the difference in this repetition-related contrast across the intentional and incidental tests, we could identify the correlates of voluntary explicit retrieval. This contrast revealed increased bilateral hippocampal responses in the intentional test, but decreased hippocampal responses in the incidental test. A similar pattern in the bilateral amygdale was further modulated by the emotionality of the word pairs, although surprisingly only in the incidental test. Parietal regions, however, showed increased repetition-related responses in both tests. These results suggest that the neural correlates of successful voluntary explicit memory differ in directionality, even if not in location, from the neural correlates of successful involuntary implicit (or explicit) memory, even when the incidental test taps conceptual processes. ■

INTRODUCTION

Human memory is multifaceted, and considerable effort has been expended in uncovering the psychological processes and brain structures that support these different facets. Two expressions of memory that have often been contrasted are explicit and implicit memory (Graf & Schacter, 1985). Explicit memories are memories that are consciously retrieved, usually implying volition in retrieval and often bringing to mind a specific episode, whereas implicit memories can influence our behavior in the absence of conscious associations with the past. Although originally intended as phenomenological labels, the terms explicit and implicit memory have also been assumed to map onto distinct neural systems (Schacter & Tulving, 1994; Schacter, 1987). Functional neuroimaging studies have since been used to investigate this claim, although often by examining each type of memory in isolation. In order to identify more precisely the brain structures that are specific to voluntary explicit

memory (and those structures that are shared), it is necessary to study explicit and implicit memory in conjunction (Voss & Paller, 2008). In particular, it is essential to compare memory tests that are matched in all respects (such as the type of memoranda, type of retrieval cue, etc.), except for their relative reliance on voluntary explicit memory.

One paradigm that has proved useful in this respect is the cued-recall/free-completion paradigm. In the "word-stem" version of this paradigm, participants are exposed to a list of words (the "study phase"), and then perform one of two tests: (1) in the intentional test, they are given the first few letters of a word ("word-stem"), and asked to recall a word from the study phase that completes that stem; (2) in the incidental test, they are given the same type of cue (a word-stem) but complete it with the first word that comes to mind (i.e., no reference is made to the previous study phase).[1] Because everything is matched except the instructions to participants, this pair of tests satisfies the "retrieval intentionality" criterion (Schacter, Bowers, & Booker, 1989). Behavioral dissociations between these two tests have been found as a function of other variables, such as the match versus mismatch in the visual/

[1]MRC Cognition & Brain Sciences Unit, Cambridge, UK, [2]University College London, UK, [3]Université de Lausanne, Switzerland

auditory modality between study and test (in some cases, even a "reversed association"; e.g., Richardson-Klavehn, Benjamin Clarke, & Gardiner, 1999).

Intentional and incidental tests such as these were originally assumed to measure explicit and implicit memory directly; in particular, the incidental test was taken as a paradigmatic example of unconscious influences from the past. However, it became increasingly clear that incidental retrieval could still be accompanied by the conscious realization that the content of that retrieval had recently been encountered (Mace, 2003; Kinoshita, 2001; Richardson-Klavehn, Gardiner, & Java, 1996). This would be an example of *involuntary* explicit memory. Another example would be *voluntary* explicit memory, that is, if participants in an incidental test spontaneously adopted an episodic retrieval strategy of thinking back to the study phase. It has therefore become important to distinguish retrieval *volition* from retrieval *awareness*, that is, voluntary/involuntary versus explicit/implicit memory (Richardson-Klavehn et al., 1996).

One way to check for "contamination" of an incidental test by voluntary explicit retrieval is to employ a manipulation that is known to affect voluntary but not involuntary retrieval. In the cued-recall/free-completion paradigm, semantic elaboration at study is such a manipulation, which increases voluntary but not involuntary retrieval (Richardson-Klavehn & Gardiner, 1998). Therefore, if an effect of this manipulation is not observed in an incidental test, one can infer that a voluntary explicit retrieval strategy was not employed. Even with this methodology, however, one cannot measure the extent to which involuntary explicit memory, that is, conscious realization of a prior exposure, occurs in an incidental test. This is particularly relevant to fMRI/PET studies because the poor temporal resolution of these techniques means that the neural correlates of such involuntary explicit retrieval can be included in neuroimaging data, even if it occurs after a completion comes to mind implicitly, that is, even if it does not affect the behavioral data (Henson, 2003). The occurrence of such involuntary explicit memory would reduce differences in the neuroimaging data from intentional versus incidental tests, and also prevent commonalities between the two tests being ascribed uniquely to, for example, implicit memory.

Thus, although there have been many neuroimaging studies that have claimed different patterns of brain activity associated with explicit and implicit memory, few used memory tests that are matched in all non-mnemonic aspects and/or provided evidence that the data from an incidental test were uncontaminated by voluntary or involuntary explicit memory. The word-stem completion paradigm is one exception, with early imaging studies (e.g., Buckner et al., 1995; Squire et al., 1992) matching the two tests according to the retrieval intentionality criterion, and later studies additionally controlling for contamination of the imaging data (e.g., by getting participants to indicate whether they had explicit memory for each completion in the incidental test; Schott et al., 2005). These studies

generally report increased hippocampal activity associated with explicit memory, but decreased occipito-temporal activity associated with implicit memory (for visual words).

However, priming (implicit memory) in the word-stem completion paradigm is believed to involve facilitation of primarily perceptual processes (at least phonological or lexical, rather than semantic, processes; Richardson-Klavehn & Gardiner, 1998; Weldon, 1991). Thus, comparison of intentional and incidental versions of the word-stem paradigm may reveal differences other than explicit memory, for instance, differences in the amount of semantic processing at test. A closer comparison is therefore achieved by using tests of conceptual priming. Conceptual priming is, like explicit memory, observed even when there is no perceptual overlap between the retrieval cue in the test phase and the target presented in the study phase (Roediger, Weldon, & Challis, 1989). By comparing an intentional test with an incidental test of conceptual memory, the brain structures that mediate voluntary episodic retrieval can be isolated more clearly, by virtue of removing shared processes relating to semantic processing.

Neuroimaging studies of conceptual priming have focused on reduced hemodynamic responses during incidental retrieval, which are often considered to reflect more efficient processing of repeated relative to novel material (Henson, 2003). Such reduced responses associated with conceptual processing have most often been found in the left inferior frontal gyrus, Brodmann's area (BA) 45/46/47 (Voss, Reber, Mesulam, Parrish, & Paller, 2008; Donaldson, Petersen, & Buckner, 2001; Buckner, Koutstaal, Schacter, & Rosen, 2000; Wagner, Koutstaal, Maril, Schacter, & Buckner, 2000; Blaxton et al., 1996; Demb et al., 1995). Only a few studies, however (Voss et al., 2008; Donaldson et al., 2001), have made direct comparisons with tests of intentional retrieval, and in these, the intentional and incidental versions of the tests differed on more dimensions than just retrieval instructions, thus failing the retrieval intentionality criterion. One exception is an early PET study (Blaxton et al., 1996) that compared cued-recall with free-completion of semantically related word pairs, and that found decreased activity in left prefrontal and left temporal cortices associated with memory retrieval in both tests (and the locations of these decreases were different from the occipital decreases associated with memory performance on two "more perceptual" tasks of cued-recall and free-completion of word fragments). This study therefore questioned the distinction between explicit and implicit memory, favoring a distinction between conceptual versus perceptual tasks. However, this study did not contrast the intentional and incidental tests directly, nor did it check for contamination of the incidental test data by voluntary explicit retrieval.

The present event-related fMRI study addresses these limitations of previous neuroimaging studies by using a cued-recall/free-completion paradigm for word associates, in which (1) the intentional (cued-recall) and incidental (free-completion) tasks were thoroughly matched, and (2)

an orthogonal manipulation of the "depth" of processing was used to check for contamination of the incidental test by explicit memory. Note that although conceptual implicit memory is also generally affected by a manipulation of depth-of-processing, in the case where the word pairs in a free association test are pre-experimentally related (as here), an effect of this variable has not been observed (i.e., a clear depth-of-processing effect has been found in intentional tests, but not for the same material in incidental tests; Ramponi, Handelsman, & Barnard, 2010; Ramponi, Nayagam, & Barnard, 2009; Ramponi, Richardson-Klavehn, & Gardiner, 2004, 2007; Vaidya et al., 1997). Unlike most previous imaging studies, the two tests were run only once, in different groups of participants. This additional precaution further reduces the risk of the contamination of the incidental test by voluntary retrieval strategies (that might occur if the same participant previously attempted an intentional test), and also helps to ensure that memory encoding in the study phase is truly incidental (i.e., participants had no warning that their memory for the word pairs would later be tested, thus were unlikely to engage in additional elaborative strategies).

Finally, we also manipulated the emotionality of the word pairs, with one half being rated relatively more emotional (including both positive and negative valence) than another half (henceforth, the "neutral" items). Explicit memory for emotional events has been shown to be superior to that for events of equivalent nature that are not considered emotional; the *memory-enhancing effect* of emotion (for reviews, see Buchanan, 2007; LaBar & Cabeza, 2006). Less is known however about the influence of emotion on implicit memory (LaBar & Cabeza, 2006). In a previous behavioral study (Ramponi et al., 2010) using the same paradigm applied here, we showed that the memory-enhancing effect of emotion was present in the intentional test, but not in the incidental test, suggesting that the emotional status of the word pairs does not affect implicit memory (or involuntary explicit memory as what might have occurred in the incidental test). Importantly, the neutral word pairs were drawn from a semantic category that was comparable to the emotion category in order to control for the semantic cohesiveness inherent to emotional items (see Talmi & Moscovitch, 2004).

In summary, the present fMRI experiment involved two groups of participants who first studied pairs of words with a preexisting semantic association, presented in two blocks. For one block, participants compared each word within the pair according to a semantic dimension (semantic condition); in the other block, the two words were compared according to a phonological dimension (phonemic condition; see Figure 1). One half of the words within each block were neutral (e.g., *ELECTRIC–LIGHT*) and the other half were emotional (e.g., *TRAGIC–DEATH*). The groups differed only in the instructions given in the subsequent test phase, in which a single "cue" word was presented and one group was told to use this cue to recall its associate from the study phase (intentional group), whereas the other group was told to produce the first word that came to mind in response to the cue (incidental group). Some of the

cues corresponded to studied words (e.g., *ELECTRIC–?, TRAGIC–?*); others were unstudied (half neutral, half emotional, e.g., *CUPBOARD–?; HOSTILE–?*). Note that the stimuli were rotated across studied/unstudied (and semantic/phonemic) conditions across participants, such that unstudied cues could also be associated with a "target" response. This allowed us to measure baseline levels of performance within the incidental and intentional tests (i.e., spontaneous completion rates). This was also important for our comparisons of the concurrent event-related fMRI data: By subtracting activity elicited by unstudied cues completed with targets from that elicited by studied cues completed with targets, we were able to isolate the brain activity associated specifically with successful memory retrieval, reducing individual variability, controlling for stimulus differences, and matching more closely other factors such as retrieval effort (Schacter, Alpert, Savage, Rauch, & Albert, 1996).

Using this fMRI measure obtained from subtracting the mean event-related response for unstudied target completions from that for studied target completions, our main questions concerned where in the brain this "repetition-related" activity: (1) differed across intentional and incidental tests, and (2) was common to both tests. More specifically, given that only the intentional test requires that participants recall a specific episode from the study phase, and given that many previous neuropsychological and neuroimaging studies have implicated the hippocampus in episodic retrieval (Eichenbaum, Yonelinas, & Ranganath, 2007; Davachi, 2006; Smith & Mizumori, 2006a, 2006b; Eichenbaum, 2001), we predicted that repetition-related activity in the hippocampus would be greater in the intentional test than in the incidental test. Secondly, given that previous neuropsychological and neuroimaging studies have implicated the amygdala in processing emotional material (Costafreda, Brammer, David, & Fu, 2008; Sergerie, Chochol, & Armony, 2008; Phelps & LeDoux, 2005; Zald, 2003) and specifically in the memory-enhancing effect of emotion (Buchanan, 2007; Roozendaal, Barsegyan, & Lee, 2007; LaBar & Cabeza, 2006; Dolcos, LaBar, & Cabeza, 2004, 2005; McGaugh, 2004; Phelps, 2004; Smith, Henson, Dolan, & Rugg, 2004; Cahill & McGaugh, 1998), we predicted that the amygdala would also show repetition-related activity only in the intentional test, but more so for emotional than neutral associates.

METHODS

Participants, Design, and Material

Thirty volunteers from the MRC Cognition and Brain Sciences Unit Volunteer Panel (19 women, all right-handed, aged 18–41 years, mean age = 25.13 years), with normal or corrected-to-normal vision, participated in this study. No participants had a history of neurological disease or head injury.

The pool of word associates used in this study included 102 word associates, 51 with emotional connotation and

Figure 1. Diagram describing the design of the experiment for the intentional test group and the incidental test group. All participants underwent the same study phase, with the phonemic and semantic conditions being blocked with block-order being randomly determined. Three lists of 34 associates, including 17 emotional and 17 matched neutral associates, were assigned to the semantic condition, the phonemic condition, or to act as unstudied associates for the test phase, following a 3 by 3 Latin square. The associates in each list were presented in random order. The test phase was identical for both groups, except for the retrieval instructions. All participants were presented with the 34 cues from the phonemic study condition, 34 cues from the semantic study condition, and 34 cues from the matched unstudied associates. There were also 34 filler cues (not shown) from unstudied associates that served simply to increase the number of unstudied cues in the test

Study phase

| **Phonemic Condition** *Which word has more syllables* | 17 Neutral Associates (e.g., ELECTRIC–LIGHT) 17 Emotional Associates (e.g., TRAGIC–DEATH) |
| **Semantic Condition** *Which word is more pleasant* | 17 Neutral Associates (e.g., WAITING–ROOM) 17 Emotional Associates (e.g., LONGING–DESIRE) |

Test phase

	Intentional test group (*n* = 15) *Recall the associate*	**Incidental test group (*n* = 15)** *First associate coming to mind*
Studied **PHONEMIC**	17 neutral associates (ELECTRIC– ?) *e.g., Target associate LIGHT (A)* 17 emotional associates (TRAGIC– ?) *e.g., Nontarget associate ACCIDENT*	17 neutral associates (ELECTRIC– ?) *e.g., Nontarget associate TRAIN* 17 emotional associates (TRAGIC– ?) *e.g., Target associate DEATH (E)*
Studied **SEMANTIC**	17 neutral associates (WAITING– ?) *e.g., Nontarget associate LIST* 17 emotional associates (LONGING– ?) *e.g., Target associate DESIRE (B)*	17 neutral associates (WAITING– ?) *e.g., Target associate ROOM (F)* 17 emotional associates (LONGING– ?) *e.g., Nontarget associate LOOK*
UNSTUDIED	17 neutral associates (CUPBOARD– ?) *e.g., Target associate DOOR (C)* 17 emotional associates (DRUG– ?) *e.g., Target associate ADDICT (D)*	17 neutral associates (CUPBOARD– ?) *e.g., Target associate DOOR (G)* 17 emotional associates (DRUG– ?) *e.g., Target associate ADDICT (H)*

Contrast:
Intentional test studied target - unstudied targets *e.g.*, (A–C)+(B–D) >
Incidental test studied targets – unstudied targets *e.g.*, (E–H)+(F–G)

phase, further discouraging incidental test participants from engaging in a voluntary retrieval strategy. The cues for the test phase were presented in random order. The intentional test participants were instructed to retrieve the specific word that had been paired with the cue word at study. If they could not retrieve the study associate, they were to think of the first word that came to mind. The incidental test participants were to think of the first word that came to mind in response to the cue word.
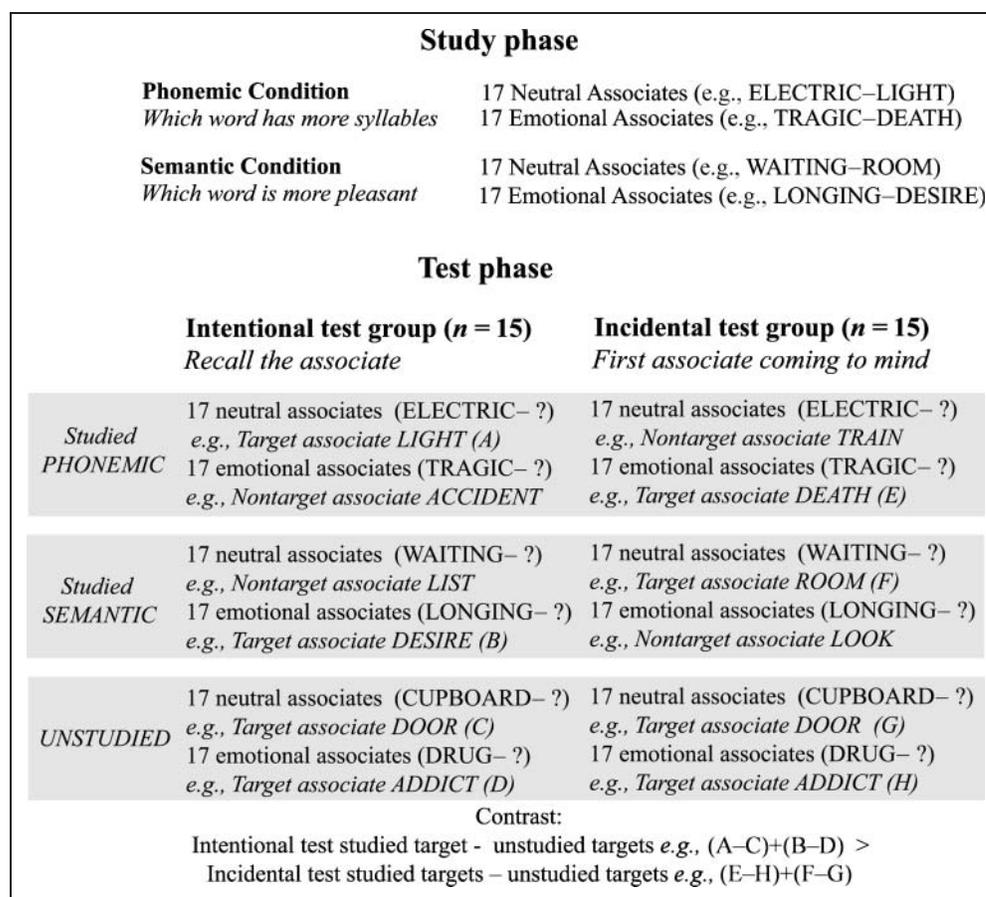
51 of neutral connotation, selected from the on-line version (www.eat.rl.ac.uk/) of the *Edinburgh Word Association Thesaurus* (Kiss, Armstrong, Milroy, & Piper, 1973). Associates were selected only if the two words constituting the pair could appeared contiguously (i.e., with no conjunction) in ordinary written and spoken language (e.g., *birthday party*). These "compound" associates were selected because they were considered to have stronger pre-existent representation (see Ramponi et al., 2007). Target words were the first or second most frequent associate of the cue words. The neutral associates all belonged to the semantic category of *household* items and activities in order to match the semantic-cohesiveness inherent in emotional words. This matching is important in order to infer that any increase in target completions for emotional relative to neutral cue words does not simply reflect the fact that emotionally related words occupy a smaller or denser part of semantic space; that is, that any effect of emotionality is more than a special case of semantic relatedness (Talmi & Moscovitch, 2004). Average percentage association strength was 23.9 (SD = 11.2). Association strength was matched between the emotional compound associates

(M = 23.8, SD = 10.6) and neutral compound associates [M = 24.1, SD = 11.9; $t(100)$ = 0.89, p > .05]. The emotional and neutral compound associates were also matched for cue [$t(100)$ = 0.78, p > .05] and target [$t(100)$ = 0.80, p > .05] frequency (Medler & Binder, 2005).

At study, all participants carried out a phonemic and a semantic task (see Figure 1 for a description of the design). Study tasks were blocked and block order was randomly determined. For each study task, participants saw a list of 34 word associates with 17 emotional associates and 17 neutral associates presented randomly. Thus, participants studied a total of 68 word associates. At test, half of the participants were randomly assigned to the intentional test and the other to the incidental test of word association. All participants were presented with all the cue words of the associates that had been presented in the study phase (phonemic/emotional, phonemic/neutral, semantic/emotional, semantic/neutral) and new cues from 34 matched associates (17 emotional and 17 neutral) not presented at study (unstudied/emotional, unstudied/neutral). The studied/unstudied status of the word pairs was rotated following a 3 by 3 Latin square (phonemic studied, semantic studied, unstudied).

Trials in the study and test lists were presented in different pseudorandom orders, with no more than three emotional or neutral associates appearing in sequence. At test, an additional 34 filler trials were presented. These trials were identical to the critical test trials, except that the cues came from other unseen compound associates of general neutral valence and similar frequency as the test associates. The filler trials were randomly intermixed with the critical trials. The purpose of these filler cues was to bring the ratio of studied to unstudied cues to 50:50, in order to reduce the chance of participants in the incidental group noticing that most cue words were from the study phase, that is, to reduce contamination of the incidental test by voluntary retrieval. The filler trials were disregarded in the analyses.

Procedure

Study Phase

Instructions and stimuli were presented using Eprime. An LCD projector was used to rear-project the stimuli onto an angled mirror placed above the participant's head. In the study phase, participants were shown instructions for the phonemic (*which word had more syllables?*) or semantic (*which word was more pleasant?*) tasks projected on the screen. Participants then practiced the task with four associates before attempting each task. The words that formed the paired associate were presented in the center of the screen in Courier New font; they were adjacent to one another and were presented for 3000 msec with a 1500-msec ISI. Participants responded by pressing with their right hand one of two buttons.

Test Phase

The test phase was preceded by a 10-min task where a surname had to be generated to a first-name cue. A first name was projected on the mirror and participants were to think of the first surname that came to mind. They then waited for a prompt on the screen, "???," that appeared for 2000 msec preceded by a cross for 500 msec, before saying into the microphone the surname that had come to mind. This procedure was identical to that of the subsequent word association test, so served to prepare the participants for the critical, subsequent word association test. It also meant that the word association test in the incidental group could be portrayed as just another free association task, in an attempt to further minimize the chance of contamination by voluntary retrieval strategies.

In the critical test, cue words were presented every 20 sec; they stayed on the screen for a random interval between 10 and 18 sec. After this interval, the word was replaced by a fixation cross, shortly replaced after 500 msec by the "???" prompt for participants to verbalize the word retrieved in response to the cue word (see below). The prompt remained for 2000 msec, again replaced by the fixation cross which stayed until the remaining period of the 20-sec in-

tertrial interval had elapsed. Then the next cue word (i.e., trial) appeared. The procedure for intentional and incidental groups was identical apart from the retrieval instructions. In the incidental test, participants had to think of the first word that came to mind in response to the cue word. In the intentional test, participants were instructed to retrieve the specific word that had been paired with the cue word at study. If they could not retrieve the study associate, they were to think of the first word that came to mind. For both tests, participants were instructed to press a button as soon as a response came to mind. However, there were told to wait until the "???" prompt before verbalizing that response. The purpose of this randomly varying delay in overt responding was to decorrelate the fMRI signal evoked by the cue word (and associated with memory retrieval) from the fMRI noise associated with movement-related artifacts caused by the speech act (Henson & Josephs, 2002).

Imaging Parameters

Thirty-two T2*-weighted transverse slices (64×64 3 mm $\times$ 3 mm pixels, TE = 30 msec, flip angle = 78°) per volume were taken using echo-planar imaging on a 3-T TIM Trio System (Siemens, Erlangen, Germany). Slices were 3-mm thick with a 0.75-mm gap, tilted approximately 30° at the front to minimize eye-ghosting, and acquired in descending order. The slices covered the whole neocortex (but sometimes excluded anterior, inferior parts of the cerebellum). The study phase was split into two sessions of 86 volumes (see Procedure); the test phase was split into three (equivalent) parts of 449, 449, and 489 volumes. The first six volumes of each session were discarded to allow for equilibrium effects. The volume repetition time (TR) was 2000 msec. The data from the study phase are not reported here.

A T1-weighted structural volume was also acquired for each participant with $1 \times 1 \times 1$ mm voxels using MP-RAGE and GRAPPA parallel imaging (flip angle = 9°; TE = 2.00 sec; acceleration factor = 2), for the purposes of anatomic localization and coregistration across participants.

A Noise Canceling Optical Microphone System for MRI (FOMRI[tm] II, Dual Channel System Optoacoustics) was used to record participants' spoken responses, which were then scored off-line. Note that noise cancellation enabled participants to speak relatively quietly (and still be intelligible against the background scanner noise), thus further reducing speech-related movement artifacts in the fMRI data. The total time each participant spent in the scanner was 1 hr 15 min. Participants were not scanned during the distracter task.

Image Analysis

Data were analyzed using the SPM5 software (www.fil. ion.ucl.ac.uk/spm5.html). Standard spatial preprocessing

was performed, including realignment to correct for head movement, rigid-body coregistration with the individual structural image, nonlinear normalization to the MNI T1-weighted template (3-mm isotropic voxels), and finally, smoothing with a Gaussian kernel of 10 mm FWHM (final estimated smoothness was approximately 15 mm FWHM isotropic).

In a first-level analysis of each participant, an event-related general linear model (GLM) was implemented by convolution of the onsets of the visual cue words with a canonical hemodynamic response function. Twelve different types of event were of interest (within both groups of participants), formed by the combination of three types of cue status (studied in phonemic task, studied in semantic task, versus unstudied), two types of cue word (emotional vs. neutral), and two types of responses (target vs. nontarget completions). The filler trials were modeled similarly, as a 13th event type of no interest. Two final event types were modeled: one locked to the onset of the speech cue and the other to the onset of the button press (because these onsets were defined regardless of the above 12 event types, they remove sources of noise without capturing much of the signal variance associated with differences between the conditions of interest). Voxelwise parameter estimates for these regressors were obtained by restricted maximum likelihood (ReML) estimation, using a temporal high-pass filter (cutoff 128 sec) to remove low-frequency drifts, and modeling temporal autocorrelation across scans with an AR(1) process.

Images of contrasts of the resulting parameter estimates (collapsed across left/right keypress) comprised the data for a second-level GLM, which treated participants as a random effect. There were eight contrasts of interest, reflecting factorial combination of intentional versus incidental test group (between-participants), semantic versus phonemic encoding, and emotional versus neutral word pairs (within-participant). These contrasts were restricted to target completions. Critically, to control for differences in activation related to chance retrieval, mirroring the behavioral analyses, each contrast subtracted from studied target trials the corresponding baseline trials, that is, where the target associate was spontaneously produced to an unstudied cue. In other words, there were four contrasts per participant: (1) phonemic emotional studied target minus emotional unstudied target; (2) semantic emotional studied target minus emotional unstudied target; (3) phonemic neutral studied target minus neutral unstudied target; and (4) semantic neutral studied target minus neutral unstudied target. Each "retrieval" contrast would therefore identify brain activity related to the effects of study: either due to successful episodic retrieval of the study associate in the intentional group, or priming of retrieval from semantic memory in the incidental group (see Introduction).

Using this second-level GLM, statistical parametric maps (SPMs) were created of the *T* or *F*-statistic, using a single pooled error estimate for all contrasts, whose nonsphericity was estimated using ReML as described in Friston et al.

(2002). The main second-level contrasts of interest, reflecting the purpose of the study, were to identify "common" and "differential" brain activity across test groups. In the case of differential activity, for example, the contrast identified where the "retrieval"-related activity owing to prior study (i.e., corrected for baseline) differed for intentional and incidental groups. To identify common activity across test groups, the main effect of study (i.e., repetition-related activity) was exclusively masked by the above interaction between the study and test groups (i.e., differential repetition-related activity), as in the original definition of a "cognitive conjunction" (Price & Friston, 1997). The SPMs for these contrasts were thresholded for at least 5 contiguous voxels that survived $p < .05$, corrected for multiple comparisons at the peak level using random field theory, either across the whole brain or within anatomical regions of interest (ROIs). Given the hypothesis outlined in the Introduction, these anatomical ROIs were defined by binary mask images containing (1) bilateral hippocampi and (2) bilateral amygdalae from the Automatic Anatomical Labeling (AAL) database (Tzourio-Mazoyer et al., 2002). Further orthogonal contrasts were then performed within the maxima that survived correction in order to investigate subsidiary effects of study task and stimulus emotionality.

RESULTS

Behavioral Results

Correct responses were responses to the cues of studied and unstudied associates that were lexically identical (but could be plurals) to the target words of the 102 selected associates (see Methods). Target completion of cues from unstudied associates estimated the probability that a word associate could be produced by chance, that is, from free association in the absence of any prior study. This formed the *baseline* rate for each participant. The mean proportions of target responses to studied and unstudied cues are reported in Table 1.

A preliminary analysis was performed to check that retrieval in the intentional and incidental tests was above baseline rates. A 2 × 2 mixed ANOVA with factors study status (studied associates vs. unstudied associates) and test group (intentional vs. incidental group) showed a significant main effect of study status [$F(1, 28) = 139.57$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .83$], with subsequent one-tailed tests indicating the expected presence of voluntary retrieval in the intentional group [$t(14) = 11.07$, $p < .001$] and involuntary retrieval (priming) in the incidental group [$t(14) = 5.72$, $p < .001$]. The main effect of test group was also significant [$F(1, 28) = 14.51$, $MSE = 0.01$, $p = .001$, $\eta_p^2 = .34$], but so was the interaction [$F(1, 28) = 13.04$, $MSE = 0.01$, $p = .001$, $\eta_p^2 = .32$], indicating that the difference between target retrieval of studied and unstudied associates was larger for the intentional group than for the incidental group, as expected. Importantly, subsequent two-tailed *t* tests showed no evidence for a difference in baseline rates (i.e., for

Table 1. Mean Proportions (and *SD*) of Phonemically and Semantically Studied Target Associates and Unstudied Target Associates (Baseline) as a Function of Associate Type (Emotional vs. Neutral) and Test Group (Intentional vs. Incidental)

	Phonemic		Semantic		Unstudied	
	M	SD	M	SD	M	SD
Intentional Test Participants (n = 15)						
Neutral	0.58	0.16	0.71	0.13	0.35	0.10
Emotional	0.60	0.15	0.76	0.14	0.30	0.09
Incidental Test Participants (n = 15)						
Neutral	0.46	0.18	0.53	0.14	0.31	0.12
Emotional	0.42	0.17	0.48	0.19	0.28	0.09

unstudied associates) between the two groups [intentional group: $M = 0.32$, $SD = 0.07$; incidental group: $M = 0.29$, $SD = 0.09$; $t(28) = 0.94$, $p = .35$], whereas retrieval of studied associates did clearly differ [intentional group: $M = 0.66$, $SD = 0.11$; incidental group: $M = 0.47$, $SD = 0.12$; $t(28) = 4.52$, $p < .001$].

We then looked at the two groups separately, taking into consideration only the responses corrected for chance retrieval, that is, with the baseline subtracted for each participant (as in Ramponi et al., 2007). These "corrected" responses are shown in Figure 2. For the intentional group, a 2 × 2 mixed ANOVA with factors of depth-of-processing (phonemic vs. semantic) and associate type (emotional vs. neutral) showed the predicted main effect of depth-of-processing [$F(1, 14) = 10.94$, $MSE = 0.03$, $p < .005$, $\eta_p^2 = .44$], with more semantically studied associates being recalled than phonemically studied associates. Also as predicted, there was a main effect of emotion, with more emotional associates being recalled than neutral associates [$F(1, 14) = 4.57$, $MSE = 0.02$, $p = .05$, $\eta_p^2 = .25$]. The interaction was not significant ($F < 1$).

For the incidental group, the main effect of depth-of-processing was not significant [$F(1, 14) = 2.116$, $MSE = 0.03$, $\eta_p^2 = .13$]. The presence of a depth-of-processing effect in the intentional test, and not in the incidental test, supports the conclusion that the incidental test was not contaminated by a voluntary retrieval strategy (Richardson-Klavehn & Gardiner, 1996; Toth, Reingold, & Jacoby, 1994; Schacter et al., 1989). Furthermore, the main effect of emotional versus neutral associates was not significant either [$F(1, 14) = 0.231$, $MSE = 0.02$, $\eta_p^2 = .16$], again in contrast to the intentional test (the interaction between depth-of-processing and associate type was also not significant, $F < 1$). This is consistent with the absence of emotion effects in conceptual implicit memory tests (Ramponi et al., 2010), and further supports different retrieval processes in the two tests, specifically the proposal that participants in the incidental group did not employ a voluntary retrieval strategy.

The mean response times (RTs) for the button presses that indicated when participants produced an associate are shown in Supplementary Table 1. Comparisons of raw RTs across test group are difficult to interpret, given the different task instructions (i.e., to produce the first word that comes to mind in the intentional test only after failing to recall the studied associate), producing different baseline RTs for unstudied items. Nonetheless, baseline-corrected RTs were tested in a 2 × 2 ANOVA within each test group. The intentional group showed only a main effect of depth-of-processing, with a greater reduction in RTs for semantically than phonemically studied items [$F(1, 14) = 9.05$, $MSE = 0.14$, $p < .01$, $\eta_p^2 = .39$]. The incidental group, however, showed a reliable interaction between depth-of-processing and associate type [$F(1, 14) = 7.90$, $MSE = 0.02$, $p < .05$, $\eta_p^2 = .36$], in addition to a main effect of depth-of-processing [$F(1, 14) = 8.68$, $MSE = 0.05$, $p < .05$, $\eta_p^2 = .38$]. (This contrasts with the lack of reliable effects of these variables on accuracy in this group.) This pattern actually reflected a greater reduction in RTs for phonemically than semantically studied items; an effect that itself was bigger for neutral than emotion associates (Supplementary Figure 1). The finding that depth-of-processing had opposite effects on RTs for studied versus unstudied items across the intentional versus incidental test groups further supports the use of different retrieval processes in the two tasks.

Imaging Results

The fMRI data of one participant in the intentional group could not be used because she failed to press the button when an associate came to mind. To control for differences in brain activity related to chance retrieval, mirroring the above behavioral analyses, we subtracted from the activity associated with target completions of studied cues the activity associated with baseline trials in which the target
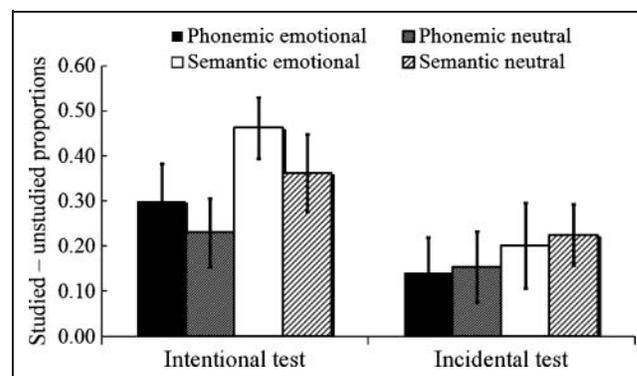
Figure 2. Mean proportions of studied target associates having subtracted baseline (spontaneous) retrieval (i.e., unstudied target associates) as a function of depth-of-processing (phonemic vs. semantic), associate type (emotional vs. neutral), and test group (intentional vs. incidental). Error bars are confidence intervals.

associate was spontaneously produced to an unstudied cue (see Methods).[2] We first tested for differences in this "repetition-related" activity between the intentional and incidental tests before investigating repetition-related activity common to both tests.

Greater Repetition-related Activity for Intentional than Incidental Tests

We predicted that retrieval in the intentional test, but not in the incidental test, requires retrieval of episodic information, that is, information about the study episode in which the cue word was paired with the target word. We therefore performed a t test in all voxels to identify brain regions where the activity associated with target retrieval from studied test cues (having subtracted activity related to spontaneous target completion from unstudied test cues) was greater in the intentional test than in the incidental test. Given prior research (see Introduction), we predicted that regions showing this activation related to episodic retrieval would include the hippocampus. We therefore used an anatomically defined mask image representing the bilateral hippocampi in order to correct for multiple statistical comparisons (see Methods).

When correcting for multiple comparisons across the whole brain, maxima were found in the right posterior hippocampus, precuneus, and left middle occipital lobes (see Table 2). Importantly, when focusing on the hippocampi, maxima were found in both left ($x = -24, y = -36, z = +3, Z = 4.34, p < .001$, corrected) and right ($x = +30, y = -39, z = 0, Z = 4.55, p < .001$, corrected) posterior hippocampi that survived correction for our bilateral anatomical ROIs. These maxima are shown in Figure 3A.

The estimated BOLD peak event-related signal changes associated with studied versus unstudied target completions for each of the four conditions in each of the two tasks are shown for these left and right hippocampal maxima in Figure 3B (activity across all 12 conditions is given separately for each group in Supplementary Tables 2 and 3). For the intentional test, it can be seen that target completion of studied cues tended to produce greater activity than target completion of unstudied cues. For the incidental test, on the other hand, completion of studied cues

produced less activity than target completion of unstudied cues. These effects are explored further in the ANOVAs below.

The same t test that was used to identify brain regions showing greater repetition-related activity in the intentional test than in the incidental test also revealed maxima in the left and right amygdale that survived correction for our anatomical ROIs ($x = -21, y = 0, z = -18, Z = 4.18, p < .01$, corrected, and $x = +24, y = -3, z = -15, Z = 3.58, p < .01$, corrected, respectively; see Figure 4A). The estimated BOLD signal changes for these maxima are shown in Figure 4B (activity across all 12 conditions separately is shown for each group in Supplementary Tables 4 and 5). These results suggest that the amygdala, like the hippocampus, plays a different role in intentional versus incidental retrieval.

Additional ANOVAs were performed for the hippocampal and amygdalar maxima in order to identify further effects of depth-of-processing and emotional/neutral associate type (analogous to the behavioral analyses). These ANOVAs also included data from both the left and right maxima, producing factors of hemisphere (left vs. right), test (intentional vs. incidental), depth-of-processing (semantic vs. phonemic), and associate type (emotional vs. neutral). Note that main effect of test was not reported because this effect was the only ANOVA effect biased by the prior selection of the hippocampal maxima by the same contrast.

For the hippocampal ANOVA, no effects reached significance [$F(1, 27) < 2.74, p > .11$; and this was true even when the ANOVAs were restricted to either the intentional or incidental test alone]. The ANOVA for the amygdala showed an interaction between test and associate type [$F(1, 27) = 5.34, MSE = 31.32, p < .05, \eta_p^2 = .17$; but no other significant effects]. A follow-up ANOVA for each test separately showed no reliable main effect or interactions in the amygdale except for a main effect of associate type in the incidental test [$F(1, 14) = 5.32, MSE = 38.94, p < .05, \eta_p^2 = .28$]. This effect represented a greater repetition-related decrease for emotional than neutral associates (Figure 4B).

To explore this unexpected amygdala effect further, post hoc tests were performed in which activity was split

Table 2. Regions That Showed Greater Event-related fMRI Response to Retrieval of Target Associates in the Intentional Relative to the Incidental Test, Having Subtracted That for Baseline (Spontaneous) Retrieval, and Correcting for Multiple Comparisons across the Whole Brain Using a Family-wise Height Threshold of $p < .05$

Cluster (Voxels)	Region (Label)	Location of Maxima (x y z, mm)	Z Statistic
63	Right hippocampus	+30 −39 0	4.55
119	Left precuneus	−15 −69 +33	4.64
137	Right precuneus	+21 −63 +33	4.45
115	Left mid-occipital	−39 −84 −24	4.82

Stereotactic coordinates correspond to the MNI template in SPM.

Figure 3. (A) Voxels in the left and right hippocampus that showed greater activity for studied relative to unstudied target completions in the intentional test group than in the incidental test group, thresholded at $p < .001$, uncorrected for purposes of display and superimposed on the mean structural image across participants (the peak voxels in both hippocampi survived correction for anatomically defined ROIs). (B) The mean estimate across participants of the peak % BOLD signal change (relative to grand mean over all voxels and scans), together with 95% confidence intervals, of the difference between studied and unstudied target completions for the hippocampal maxima in (A), further split by the depth-of-processing and the emotionality of the associates.
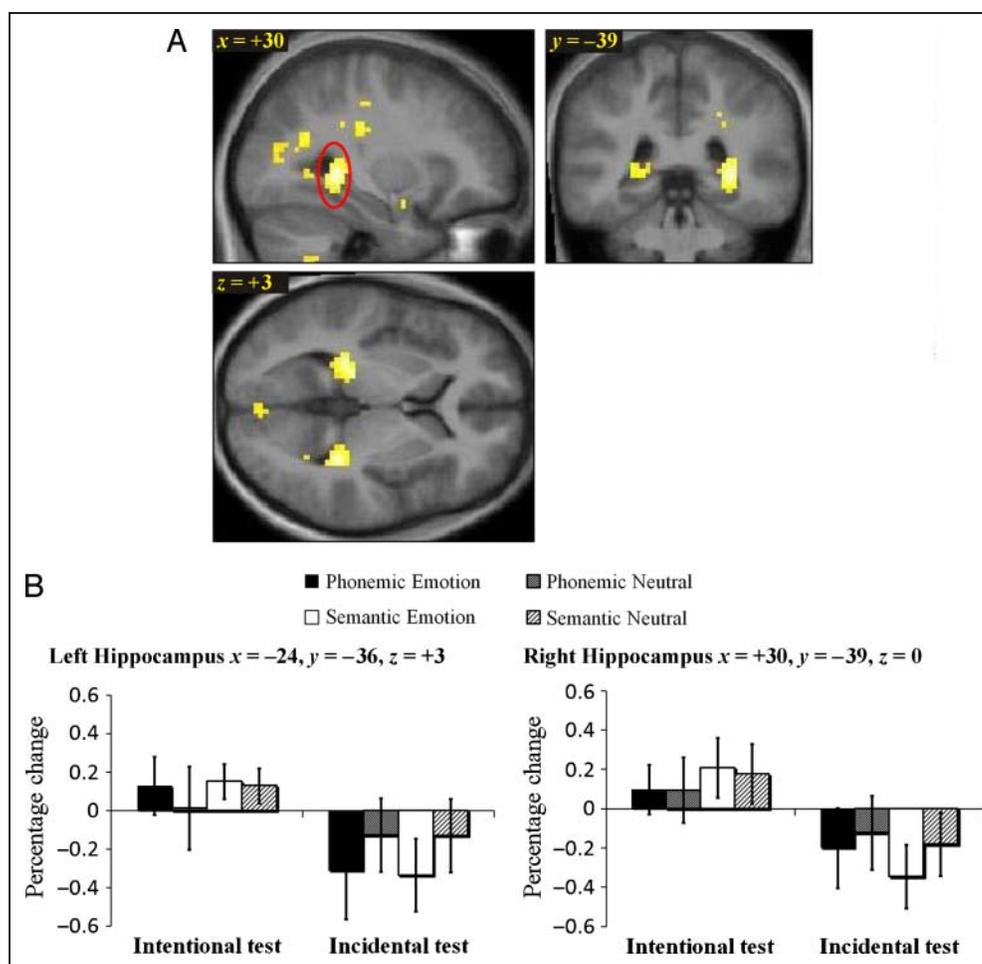
according to studied and unstudied target completions (see Supplementary Tables 4 and 5). For unstudied completions, there were no reliable effects in an ANOVA with factors of hemisphere (left vs. right amygdala maxima), test (intentional vs. incidental), and associate type (emotional vs. neutral) [$F(1, 27) < 2.86, p > .10$]. For studied completions, the same ANOVA with the additional factor of depth-of-processing (semantic vs. phonemic) showed only a reliable interaction between associate type and test [$F(1, 27) = 7.21, MSE = 8.59, p < .05, \eta_p^2 = .21$]. Subsequent t test, collapsing over hemisphere and depth-of-processing, showed greater amygdala activity for emotional than neutral studied completions in the intentional test [$t(13) = 1.98, p < .05$, one-tailed; see Supplementary Tables 4 and 5]. This is consistent with the behavioral results of increased target completions for emotional than for neutral associates in the intentional test. However, less amygdala activity was seen for emotional than for neutral completions in the incidental test [$t(14) = -2.00, p < .05$, one-tailed]. Interpretation of these results is considered further in the Discussion.

No regions showed greater repetition-related activity for target completions in the incidental test than in the intentional test (i.e., the reverse SPM contrast to that above) that survived our corrected thresholds.

Retrieval-related Activity Common to Both Intentional and Incidental Tests

We also asked which regions showed increased repetition-related activity associated with target completions that was common to both tests. To do this, we tested the t contrast for repetition-related activation averaged across intentional and incidental tests, and masked the results exclusively with the F-contrast for the above difference in repetition-related activity for intentional versus incidental tests. Such exclusive masking removes voxels in which there is evidence that the repetition-related activity differed across tasks; and by thresholding this mask contrast very liberally ($p < .05$, uncorrected), we produce a stringent test of activity common to both tasks (this is formally equivalent to a "cognitive conjunction" as originally defined by Price & Friston, 1997). This procedure identified clusters in the posterior cingulate, precuneus, bilateral intraparietal sulci, and right insula that survived correction for the whole brain (see Table 3 and Figure 5A).

As above, follow-up ANOVAs were performed on each parietal region in Figure 5 to examine effects of depth-of-processing and associate type. All regions showed a main effect of depth-of-processing ($Fs > 10.5, ps < .005$), with greater activity for semantically than phonemically studied

items, except perhaps the posterior cingulate, which just failed to reach significance ($F = 3.977$, $MSE = 9.75$, $p = .06$, $\eta_p^2 = .12$; see Figure 5B). The right intraparietal sulcus region also showed a three-way interaction between depth-of-processing, associate type, and test [$F(1, 27) = 6.28$, $MSE = 5.36$, $p < .05$, $\eta_p^2 = .19$], which was not explored further given the number of tests performed. No other effect reached significance in any region.

Finally, for completeness, we asked which regions showed decreased activity associated with target retrieval that was common to both tests (using the opposite direction of t contrast as above and the same exclusive mask). This identified two clusters, one in left medial anterior prefrontal cortex ($x = -12$, $y = +57$, $z = +12$, $Z = 5.13$) and one in the right central sulcus ($x = +33$, $y = -9$, $z = +36$, $Z = 5.61$).

DISCUSSION

In this study, we probed participants' memory with either an intentional or incidental test of word association. The two tests were identical in all respects apart from the retrieval instructions. After studying semantically related emotional and neutral word pairs, all participants were shown a single word (cue) at test, half of which had been presented at study. Participants in the intentional group were instructed to recall the associated word (target) that was paired at study, and if they could not, then to say the first associate that came to mind. Participants in the incidental group were instructed to always say the first associate that came to mind. The test cues that had not been seen at study provided a measure of baseline performance, that is, spontaneous production of target words. The main aim of the baseline-corrected contrast between the two tests was to identify the brain regions that supported voluntary retrieval from episodic memory (explicit memory), once involuntary retrieval (conceptual priming) and chance retrieval (from semantic memory) were taken into account.

It was first required to show behaviorally that the two tests tapped different retrieval processes. Both groups had similar baseline performance (target completions of unstudied words), but the intentional group retrieved more targets for studied words than did the incidental group. The intentional group also retrieved more targets studied with the semantic study task relative to those studied with the phonemic task. By contrast, this depth-of-processing effect was not reliable in the incidental group. This dissociation between the two tests supports the

Figure 4. (A) Voxels in the left and right amygdala that showed greater activity for studied relative to unstudied target completions in the intentional test group than in the incidental test group. (B) Mean repetition-related response amplitude for both amygdalae (see Figure 3 legend for more details).
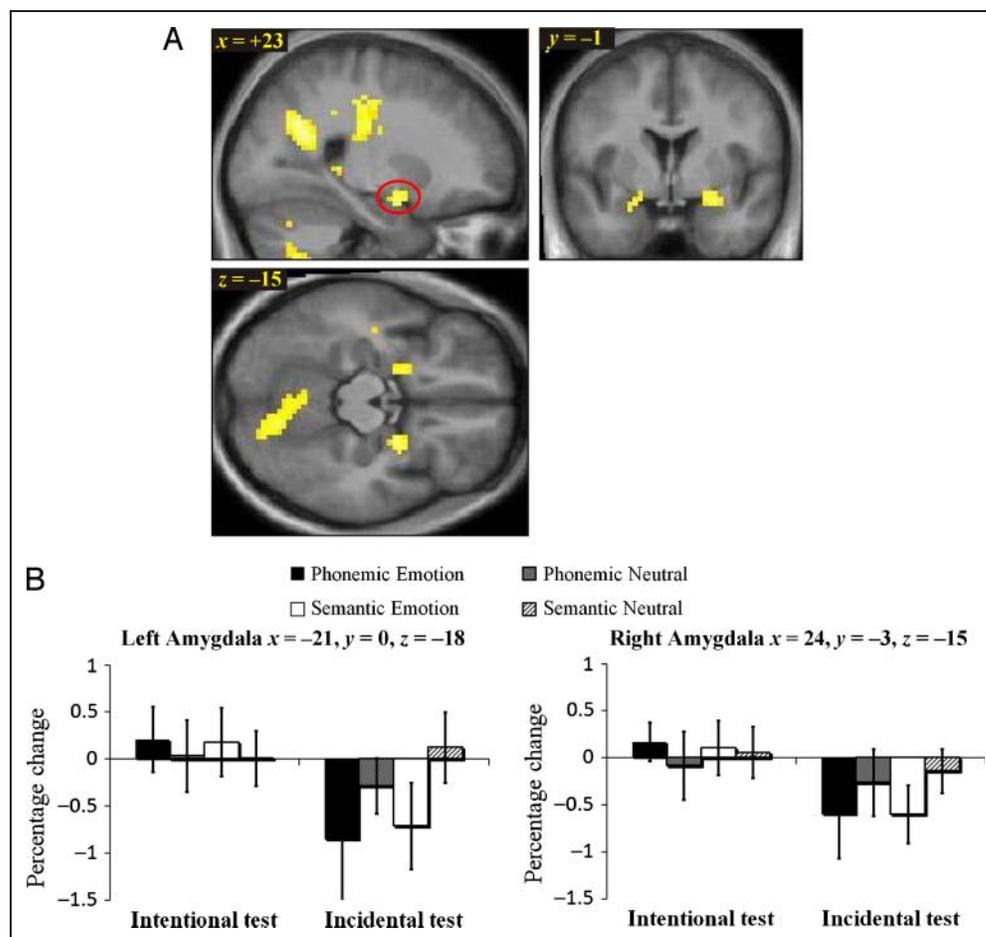
Table 3. Regions That Showed Comparable Increases in Event-related fMRI Response to Retrieval of Target Associates in the Intentional and Incidental Tests, Having Subtracted That for Baseline (Spontaneous) Retrieval, and Correcting for Multiple Comparisons across the Whole Brain Using a Family-wise Height Threshold of $p < .05$

Cluster (Voxels)	Region (Label)	Location of Maxima (x y z, mm)	Z Statistic
16	Right insula	+33 +24 +3	5.07
152	Posterior cingulate	−3 −33 +27	6.77
280	Left intraparietal sulcus	−33 −57 +39	7.20
		−48 −51 +45	4.99
92	Right intraparietal sulcus	36 −57 +42	5.30
		+39 −66 +45	5.12
		+48 −45 +42	4.91
197	Left precuneus	−12 −63 +33	7.53
		−9 −69 +42	7.99
	Right precuneus	+12 −69 +39	6.64

Stereotactic coordinates correspond to the MNI template in SPM.

contention that different retrieval processes were engaged, and that contamination of the incidental group's behavioral performance by voluntary explicit retrieval did not occur. However, this result does not preclude the occurrence of involuntary explicit memory in the incidental test.

The other key behavioral finding was that the emotionality of the word pairs also affected the two tests differently: More emotional than neutral studied targets were retrieved in the intentional test, but again, this was not the case in the incidental test. This replicates our previous results (Ramponi et al., 2010), and further supports the claim that the incidental test was not contaminated by voluntary explicit retrieval. It is also consistent with the claim that the memory-enhancing effect of emotion (for reviews, see Buchanan, 2007; LaBar & Cabeza, 2006) may represent a specific advantage in associating items to an episodic context (Ramponi et al., 2010; Hadley & MacKay, 2006).

Differential Hippocampal Activity across Tests

The critical difference between retrieval in the intentional and incidental tests was the requirement to reinstate a representation of not just the associate, but also of the episodic context in which it had been studied. Based on the extensive evidence that the hippocampus is a critical structure in such episodic binding (Eichenbaum et al., 2007; Davachi, 2006; Smith & Mizumori, 2006a; Eichenbaum, 2001), we predicted that the hippocampus would be more active for target completions in the intentional than in the incidental test. This prediction was confirmed, in parts of both left and right hippocampi, providing some of the first evidence that the hippocampus is involved specifically in associating items with their episodic context, rather than just associations between two items (at least when those items are already semantically related). It is important to note that the specific contrast that tested this prediction en-

tailed an interaction between test group and studied versus unstudied target completions, thus controlling for general differences between the two tests (e.g., retrieval effort) and for differences in the stimuli (which were rotated across studied and unstudied conditions across participants).[3]

Somewhat surprisingly, however, the hippocampal voxels revealed by the above contrast also showed a repetition-related decrease in the incidental test. Such decreases (sometimes called repetition suppression) are normally associated with priming (Henson, 2003). To our knowledge, repetition suppression effects in the hippocampus during incidental tests have not been reported before (possibly because conceptual priming has not been studied as extensively as perceptual priming with fMRI). One explanation could be that the hippocampus shows a novelty response to unstudied cues (perhaps related to new episodic encoding; Stark & Okado, 2003; Buckner, Wheeler, & Sheridan, 2001). This would explain the decreased fMRI responses for studied relative to unstudied targets in the incidental test. To explain the increased responses for studied relative to unstudied targets in the intentional test, one would have to assume that the additional hippocampal activity associated with retrieval of episodic context overcomes this novelty response. In any case, the results clearly indicate an important difference between the precise component processes engaged during our intentional and incidental tests (see later).

The hippocampal (and amygdala) repetition-related decreases in the incidental test were highlighted because they emerged from the planned comparison of repetition effects across intentional and incidental tests. Nonetheless, repetition-related decreases in the incidental test were found in other areas (see Supplementary Table 6). Surprisingly, these did not include (at our thresholds) the left inferior frontal gyrus, in which such decreases have been previously associated with conceptual priming (Voss et al., 2008; Donaldson et al., 2001; Buckner et al., 2000;

Wagner et al., 2000; Blaxton et al., 1996; Demb et al., 1995). One reason for this may be that most previous studies have assessed what has been termed *identification* priming (see Gabrieli et al., 1999), rather than the *production* priming assessed here. However, this explanation would not account for the priming-related decreases in left inferior frontal gyrus observed by Blaxton et al. (1996), who assessed production priming of word pairs as here
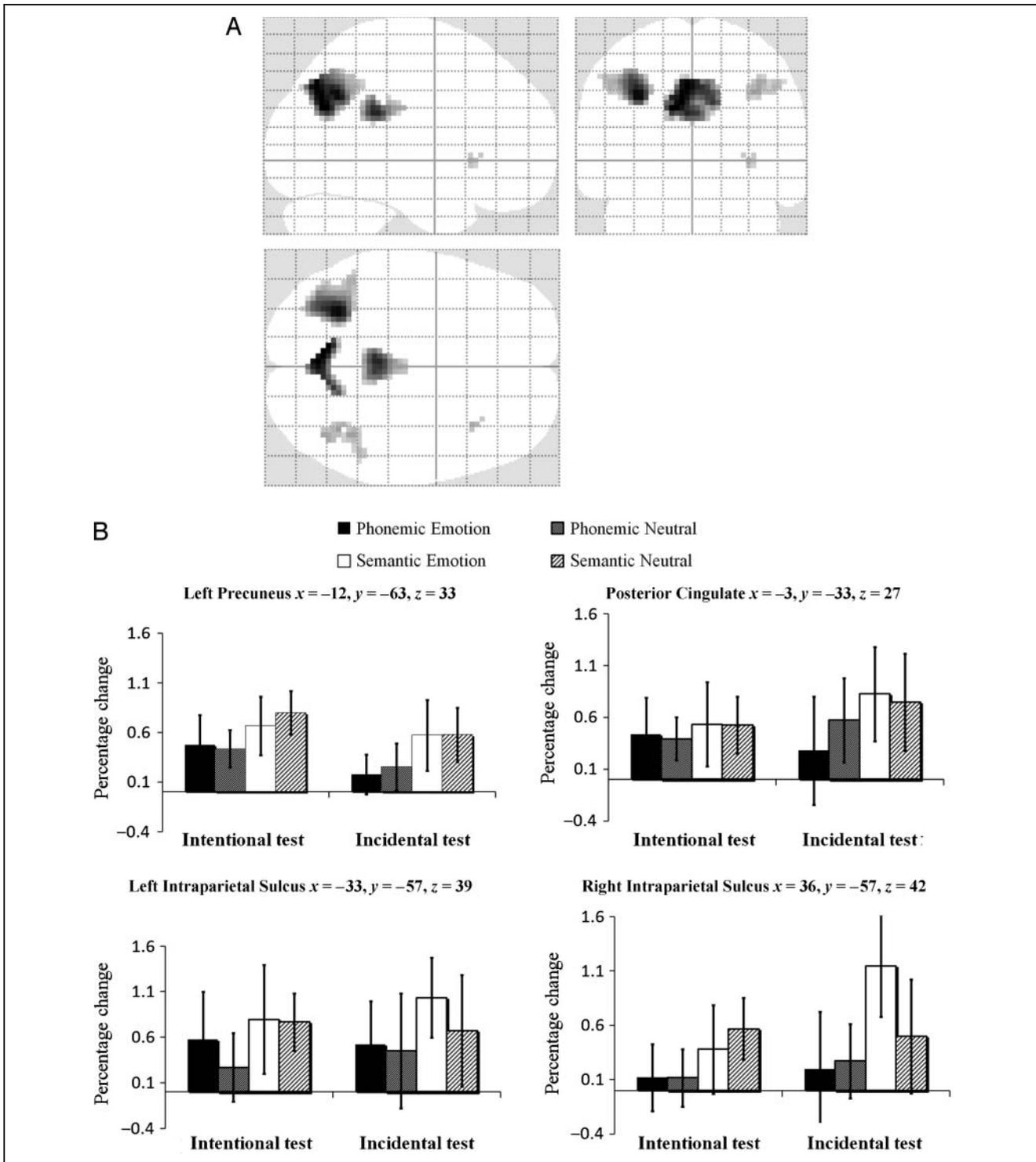
Figure 5. (A) maximal intensity projection (MIP) of the SPM of regions where the repetition-related BOLD activation was similar across the intentional and incidental tests (thresholded at $p < .05$, corrected, one-tailed, exclusively masked with the difference between intentional and incidental tests, at $p < .05$ uncorrected, two-tailed). (B) Mean repetition-related response amplitude for selected parietal regions (for more details, see Figure 3 legend).

(although indirectly via a blocked rather than event-related comparison). One difference between the current study and the Blaxton et al. study is the type of association between the words in each pair. The present study used compound associates, where the target associate is the strongest associate of the cue, whereas the Blaxton et al. study used weak associates. Thus, it is possible that our current paradigm is less sensitive to hemodynamic differences in the left inferior frontal gyrus because the generation of associates required less extensive semantic search owing to their high strength of association with the cue.

In the contrast between intentional and incidental tests, reliable differences were also observed in the bilateral precuneus. Precuneus activations are often observed in episodic retrieval (for a review, see Cavanna & Trimble, 2006). One hypothesis views the precuneus as part of a network that mediates self-awareness, hence playing a central role in the modulation of conscious processes (Cavanna & Trimble, 2006). In this vein, our precuneus result could indicate greater conscious awareness of episodic information (possibly involving oneself) in the intentional test relative to the incidental test, as might be expected.

Differential Amygdala Activity across Tests and Emotional Associates

Processing of emotional information has generally been associated with the amygdala (Costafreda et al., 2008; Sergerie et al., 2008; Phelps & LeDoux, 2005; Zald, 2003), including when that information is retrieved from episodic memory. According to one view (Dolan, Lane, Chua, & Fletcher, 2000; see also, Kensinger & Schacter, 2005; Weiskrantz, 1956), amygdala activation during episodic retrieval arises because the representations of the retrieved memories are tagged with their previous behavioral significance. According to this view, the amygdala activation does not augment retrieval of emotional information; rather, its activation is just a consequence of retrieving emotional information (and the memory-enhancing effect of emotion that we found behaviorally in the intentional but not in the incidental tests would presumably then have a different neural cause probably occurring at encoding; see McGaugh, 2004). This view would imply that amygdala activation should occur when representations of emotional items are recalled, even though their episodic context is not required, as in the present incidental test. An alternative view is that amygdala activity is an important contributor to the memory-enhancing effect of emotion (i.e., plays a causal role in our behavioral results; see Buchanan, 2007; LaBar & Cabeza, 2006; Dolcos et al., 2005; Maratos, Dolan, Morris, Henson, & Rugg, 2001).

Voxels within the left and right amygdale did show an interaction between studied versus unstudied cues and intentional versus incidental test. Indeed, they showed a pattern of repetition-related responses across the two tests that was similar to that in the hippocampus (note that this similarity is unlikely to be due to spatial smoothing, as the peak coordinates were at least 30 mm apart; nor is this pattern a simple reflection of RTs; cf. Figures 3B and 4B with Supplementary Figure 1). Furthermore, and unlike the hippocampi, the amygdalae showed an interaction between the type of test and the emotionality of the associates. To explore this interaction further, amygdala responses were split according to studied versus unstudied cues. Analysis of unstudied cues did not reveal any effects. Analysis of studied cues, on the other hand, revealed a reliable interaction, suggesting that the amygdala does not simply respond to retrieval of emotional information (from either explicit or implicit memory), but plays a more specific role in retrieval of that information from episodic memory, supporting the view that the amygdala contributes to the emotion-enhancing effect in voluntary explicit memory.

However, analysis of simple effects showed that this interaction between associate type and test group in the amygdala response was driven by both increased activity for emotional relative to neutral studied cues in the intentional test, and by decreased activity for emotional relative to neutral studied cues in the incidental test. The former increase in the intentional test is consistent with the behavioral data that showed an increased incidence of target completions for emotional than neutral cues in the intentional test. The latter decrease for studied cues explains why there was a repetition-related decrease in the "baseline-corrected" contrast for the amygdala in the incidental test (Figure 4B). However, it is more difficult to reconcile with the behavioral data, given that there was no reliable effect of emotional versus neutral cues in the incidental test.

Finally, it is important to note that the fMRI results reflect the mean activity per trial of each condition; thus, it is possible for a manipulation such as depth-of-processing or emotionality of associates to increase the number of target completions (as in the intentional test) without necessarily increasing the mean fMRI activity associated with those completions. Thus, deeper processing at study may increase the *proportion* of target completions, without affecting the average activity *per completion* within the hippocampus (and conversely, priming of emotional associates in the incidental test may produce reduced activity in the amygdala without necessarily increasing the number of such completions). This also means that the results of previous blocked designs (Buckner et al., 1995; Squire et al., 1992) cannot be easily compared with the present event-related results because the mean activity during the blocks in those studies is also affected by differences in, for example, the number of target completions across intentional and incidental tests.

Common Parietal Repetition-related Response Increases across Tests

In addition to the differences between repetition-related activity and intentional versus incidental tests in the MTL, we also found commonalities in the repetition-related responses in other cortical regions. Specifically, we found common repetition-related response increases in several

medial and lateral aspects of parietal cortex that are consistent with many previous studies of explicit memory retrieval (see Wagner, Shannon, Kahn, & Buckner, 2005, for a review). Furthermore, these repetition-related activations were generally greater when the targets had been studied semantically rather than phonologically, consistent with previous studies (Henson, Hornberger, & Rugg, 2005), and consistent with the activations being driven by episodic retrieval (Vilberg & Rugg, 2009). The question then is why similar parietal activations were found in the incidental test. Are these activations related to episodic retrieval? One hypothesis is that they are, reflecting involuntary explicit memory. As argued above, after a target word is "automatically" produced in the incidental test, participants may then realize that they had seen that associate during the previous study phase (Mace, 2003; Kinoshita, 2001; Richardson-Klavehn et al., 1996). If such involuntary memory occurred on the majority of target-completion trials in the incidental test, then this could explain why the parietal activation patterns were similar to those in the intentional test.

An alternative explanation of the parietal repetition-related activations across tests relates to the attentional account of parietal involvement in memory (Cabeza, Ciaramelli, Olson, & Moscovitch, 2008). According to this account, more dorsal regions of parietal cortex are involved in top–down attentional processes that maintain retrieval goals and regulate memory-related activity in the medial temporal lobe, whereas more ventral regions of parietal cortex signal that a change in the locus of attention ought to occur through a "bottom–up" process when "relevant" memories, retrieved by the medial temporal lobe, are detected. The authors argue that bottom–up attention can happen when memories involuntarily enter consciousness and take over attentional resources. This could explain the parietal activity (which tended to be in more ventral parietal regions identified by Cabeza et al.) that we found in the intentional test. In order to explain why this "bottom–up" retrieval and ventral parietal activity also occurs in the present incidental test, this attentional account would need to assume that attention is attracted even when the involuntary memories that enter consciousness are irrelevant to the current (task) goals.

In either case, the hypothesis of involuntary explicit retrieval raises further questions about the hippocampus: If involuntary explicit memory were the reason for the parietal activations, why did the hippocampus not show the same repetition-related response increases in the incidental test? One implication is that the hippocampus is selectively activated during voluntary retrieval from episodic memory, whereas the parietal regions echo episodic retrieval regardless of whether that retrieval was voluntary or involuntary. This interesting relationship between hippocampal activity and voluntary retrieval may be linked to recent views of the hippocampus being involved in replaying past information or relation between segments of experiences, and in *imagining* or exploring possible future scenarios (e.g., Buckner, 2010): Only when participants *explore* an (explicit) memory trace that has been retrieved, like verifying whether a recognized item appeared in the given context, is the hippocampus involved.

An alternative explanation is that the differing hippocampal and parietal activity patterns reflect explicit memory for different types of information: The hippocampal activation that only occurred in the intentional test may reflect reinstantiation of context–item associations (necessary to link the target item with the specific context of the study phase), whereas the parietal activations that occurred in both tests may reflect explicit memory for item–item associations (rather than between item and context); that is, recollection of (or familiarity for) the specific pairing of cue and target words. Another possible difference in the content of episodic retrieval is that the parietal activations reflect retrieval of egocentric contextual representations, that is, of the location of stimuli relative to oneself, whereas the hippocampal activity reflects constraints on retrieval of allocentric representations that are only imposed in the present intentional test (Bird & Burgess, 2008). Whatever the explanation, this striking dissociation between hippocampal and parietal activity is the first neuroimaging dissociation between retrieval-related activity in these two regions to be reported; a dissociation that may help tease apart the functions of these regions in relation to explicit memory retrieval, voluntary or involuntary, in future investigations.

Conclusion

This study compared two tests that were identical in all respects apart from the retrieval instructions. Such a "tight" comparison is better able to identify the structures crucial for supporting explicit memory processes than is possible when explicit and implicit memory are studied in isolation (see also Voss & Paller, 2008, for a similar argument). There was compelling evidence that the hippocampus mediates voluntary retrieval (in the intentional group), once involuntary retrieval (in the incidental group) and spontaneous retrieval from semantic memory (the baseline) have been subtracted. This hippocampal response pattern was unlike that found in parietal regions, which showed similar levels of repetition-related activation in both intentional and incidental groups. This could reflect involuntary explicit retrieval in the incidental test, although the extent of involuntary explicit memory could not be measured in the current design. Together, these fMRI data represent an important dissociation between the functions of MTL and parietal cortex during episodic retrieval. Furthermore, the amygdala showed a response profile similar to that of the hippocampus, although with evidence of a modulation by the emotionality of the word pairs. This modulation appeared to be driven by a greater response to emotional than neutral studied cues in the intentional test, consistent with the behavioral data, and supporting a contributory role for the amygdala in the memory-enhancing effect of

emotion. However, there was also a reduced amygdala response for emotional than neutral cues in the incidental test, which is less easily explained. Together, these results emphasize the need for further, more detailed investigation of the precise component processes that are engaged during tests of implicit versus explicit memory, even if those tests superficially appear to differ minimally, for instance, in terms of retrieval instructions.

Acknowledgments

Notes

1. These are also called "direct" versus "indirect" tests, respectively (Richardson-Klavehn & Bjork, 1988), although we use the labels "intentional" and "incidental" here as they more intuitively capture our design. More generally, it is important to distinguish the type of test (intentional/incidental), the phenomenological experience (explicit/implicit), and the participant's strategy (voluntary vs. involuntary episodic retrieval).

2. A comparison of the fMRI activity associated with the baseline trials between the two tests showed no reliable difference (as is also the case in the behavioral data), suggesting that participants produced unstudied targets in a similar manner in both tests; that is, significant differences in retrieval intention were not detected when the two groups produced a target associate to unstudied cues, suggesting that in the intentional test, participants freely associated when cues (from unstudied pairs) were not recognized as studied.

3. A comparison between the intentional and incidental tests during the production of a studied nontarget (i.e., nonsuccessful retrieval) was performed to test whether differences in the hippocampus relate to retrieval mode rather than retrieval success. In this contrast, no reliable differences between tests survived correction for the bilateral anatomical ROI. This is consistent with the claim that hippocampal involvement is critical for the successful association of items with their episodic context.

REFERENCES

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: Insights from spatial processing. *Nature Reviews Neuroscience, 9,* 182–194.

Blaxton, T. A., Bookheimer, S. Y., Zeffiro, T. A., Figlozzi, C. M., Gaillard, W. D., & Theodore, W. H. (1996). Functional mapping of human memory using PET: Comparisons of conceptual and perceptual tasks. *Canadian Journal of Experimental Psychology—Revue Canadienne De Psychologie Experimentale, 50,* 42–56.

Buchanan, T. W. (2007). Retrieval of emotional memories. *Psychological Bulletin, 133,* 761–779.

Buckner, R. L. (2010). The role of the hippocampus in prediction and imagination. *Annual Review of Psychology, 61,* 27–48.

Buckner, R. L., Koutstaal, W., Schacter, D. L., & Rosen, B. R. (2000). Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain, 123,* 620–640.

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R., & Raichle, M. E. (1995). Functional anatomical studies of explicit and implicit memory retrieval tasks. *Journal of Neuroscience, 15,* 12–29.

Buckner, R. L., Wheeler, M. E., & Sheridan, M. A. (2001). Encoding processes during retrieval tasks. *Journal of Cognitive Neuroscience, 13,* 406–415.

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: An attentional account. *Nature Reviews Neuroscience, 9,* 613–625.

Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences, 21,* 294–299.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain, 129,* 564–583.

Costafreda, S. G., Brammer, M. J., David, A. S., & Fu, C. H. Y. (2008). Predictors of amygdala activation during the processing of emotional stimuli: A meta-analysis of 385 PET and fMRI studies. *Brain Research Reviews, 58,* 57–70.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology, 16,* 693–700.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., & Gabrieli, J. D. E. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *Journal of Neuroscience, 15,* 5870–5878.

Dolan, R. S., Lane, R., Chua, P., & Fletcher, P. (2000). Dissociable temporal lobe activations during emotional episodic memory retrieval. *Neuroimage, 11,* 203–209.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron, 42,* 855–863.

Dolcos, F., Labar, K. S., & Cabeza, R. (2005). Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 2626–2631.

Donaldson, D. I., Petersen, S. E., & Buckner, R. L. (2001). Dissociating memory retrieval processes using fMRI: Evidence that priming does not support recognition memory. *Neuron, 31,* 1047–1059.

Eichenbaum, H. (2001). The hippocampus and declarative memory: Cognitive mechanisms and neural codes. *Behavioural Brain Research, 127,* 199–207.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30,* 123–152.

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

Gabrieli, J. D. E., Stone, M., Thompson-Schill, S. L., Tinklenberg, J. R., Yesavage, J. A., Vaidya, C. J., et al. (1999). Convergent behavioral and neuropsychological evidence for a distinction between identification and production forms of repetition priming. *Journal of Experimental Psychology: General, 128,* 479–498.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 501–518.

Hadley, C. B., & MacKay, D. G. (2006). Does emotion help or hinder immediate memory? Arousal versus priority-binding mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 79–88.

Henson, R. N. A. (2003). Neuroimaging studies of priming. *Progress in Neurobiology, 70,* 53–81.

Henson, R. N. A., Hornberger, M., & Rugg, M. D. (2005). Further dissociating the processes involved in recognition memory: An fMRI study. *Journal of Cognitive Neuroscience, 17,* 1058–1073.

Henson, R. N. A., & Josephs, O. (2002). *Spoken cued-recall during event-related fMRI.* Paper presented at the Human Brain Mapping Meeting, Sendai, Japan.

Kensinger, E. A., & Schacter, D. L. (2005). Retrieving accurate and distorted memories: Neuroimaging evidence for effects of emotion. *Neuroimage, 27,* 167–177.

Kinoshita, S. (2001). The role of involuntary aware memory in the implicit stem and fragment completion tasks: A selective review. *Psychonomic Bulletin & Review, 8,* 58–69.

Kiss, G. R., Armstrong, G., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken, R. W. Bailey, & N. Hamilton-Smith (Eds.), *The computer and literary studies* (pp. 153–165). Edinburgh: University Press.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7,* 54–64.

Mace, J. H. (2003). Involuntary aware memory enhances priming on a conceptual implicit memory task. *American Journal of Psychology, 116,* 281–290.

Maratos, E. J., Dolan, R. J., Morris, J. S., Henson, R. N. A., & Rugg, M. D. (2001). Neural activity associated with episodic memory for emotional context. *Neuropsychologia, 39,* 910–920.

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience, 27,* 1–28.

Medler, D. A., & Binder, J. R. (2005). *MCWord: An on-line orthographic database of the English language.* Retrieved May 2008 from www.neuro.mcw.edu/mcword/.

Phelps, E. A. (2004). Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology, 14,* 198–202.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron, 48,* 175–187.

Price, C. J., & Friston, K. J. (1997). Cognitive conjunction: A new approach to brain activation experiments. *Neuroimage, 5,* 261–270.

Ramponi, C., Handelsman, G., & Barnard, P. J. (2010). The memory enhancement effect of emotion is absent in conceptual implicit memory. *Emotion, 10,* 294–299.

Ramponi, C., Nayagam, J. S., & Barnard, P. J. (2009). Conceptual implicit memory in subclinical depression. *Cognition & Emotion, 23,* 551–568.

Ramponi, C., Richardson-Klavehn, A., & Gardiner, J. M. (2004). Level of processing and age affect involuntary conceptual priming of weak but not strong associates. *Experimental Psychology, 51,* 159–164.

Ramponi, C., Richardson-Klavehn, A., & Gardiner, J. M. (2007). Component processes of conceptual priming and associative cued recall: The roles of preexisting representation and depth of processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 843–862.

Richardson-Klavehn, A., Benjamin Clarke, A. J., & Gardiner, J. M. (1999). Conjoint dissociations reveal involuntary "perceptual" priming from generating at study. *Consciousness and Cognition, 8,* 271–284.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology, 39,* 475–543.

Richardson-Klavehn, A., & Gardiner, J. M. (1996). Cross-modality priming in stem completion reflects conscious memory, but not voluntary memory. *Psychonomic Bulletin & Review, 3,* 238–244.

Richardson-Klavehn, A., & Gardiner, J. M. (1998). Depth-of-processing effects on priming in stem completion: Tests of the voluntary-contamination, conceptual-processing, and lexical-processing hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 593–609.

Richardson-Klavehn, A., Gardiner, J. M., & Java, R. I. (1996). Memory: Task dissociations, process dissociations and dissociations of consciousness. In G. Underwood (Ed.), *Implicit cognition* (pp. 85–158). Oxford: Oxford University Press.

Roediger, H. L., Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3–14). Hillsdale, NJ: Erlbaum.

Roozendaal, B., Barsegyan, A., & Lee, S. (2007). Adrenal stress hormones, amygdala activation, and memory for emotionally arousing experiences. In *Progress in brain research* (Vol. 167, pp. 79–97). Amsterdam: Elsevier.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 501–518.

Schacter, D. L., Alpert, N. M., Savage, C. R., Rauch, S. L., & Albert, M. S. (1996). Conscious recollection and the human hippocampal formation: Evidence from positron emission tomography. *Proceedings of the National Academy of Sciences, U.S.A., 93,* 321–325.

Schacter, D. L., Bowers, J., & Booker, J. (1989). Intention, awareness, and implicit memory: The retrieval intentionality criterion. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory* (pp. 47–65). Hillsdale, NJ: Erlbaum.

Schacter, D. L., & Tulving, E. (Eds.) (1994). *Memory systems 1994.* Cambridge, MA: MIT Press.

Schott, B. H., Henson, R. N., Richardson-Klavehn, A., Becker, C., Thoma, V., Heinze, H. J., et al. (2005). Redefining implicit and explicit memory: The functional neuroanatomy of priming, remembering, and control of retrieval. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 1257–1262.

Sergerie, K., Chochol, C., & Armony, J. L. (2008). The role of the amygdala in emotional processing: A quantitative meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 32,* 811–830.

Smith, A. P. R., Henson, R. N. A., Dolan, R. J., & Rugg, M. D. (2004). fMRI correlates of the episodic retrieval of emotional contexts. *Neuroimage, 22,* 868–878.

Smith, D. M., & Mizumori, S. J. Y. (2006a). Hippocampal place cells, context, and episodic memory. *Hippocampus, 16,* 716–729.

Smith, D. M., & Mizumori, S. J. Y. (2006b). Learning-related development of context-specific neuronal responses to places and events: The hippocampal role in context processing. *Journal of Neuroscience, 26,* 3154–3163.

Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O., & Raichle, M. E. (1992). Activation of the hippocampus in normal humans: A functional anatomical study of memory. *Proceedings of the National Academy of Sciences, U.S.A., 89,* 1837–1841.

Stark, C. E. L., & Okado, Y. (2003). Making memories without trying: Medial temporal lobe activity associated with incidental memory formation during recognition. *Journal of Neuroscience, 23,* 6748–6753.

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition, 32,* 742–751.

Toth, J. P., Reingold, E. M., & Jacoby, L. L. (1994). Toward a redefinition of implicit memory: Process dissociations following elaborative processing and self-generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 290–303.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15,* 273–289.

Vaidya, C. J., Keane, M. M., Gutierrez-Rivas, H., Gabrieli, J. D. E., Monti, L. A., & Zarella, M. M. (1997). Evidence for multiple mechanisms of conceptual priming on implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1324–1343.

Vilberg, K. L., & Rugg, M. D. (2009). An investigation of the effects of relative probability of old and new test items on the neural correlates of successful and unsuccessful source memory. *Neuroimage, 45,* 562–571.

Voss, J. L., & Paller, K. A. (2008). Brain substrates of implicit and explicit memory: The importance of concurrently acquired neural signals of both memory types. *Neuropsychologia, 46,* 3021–3029.

Voss, J. L., Reber, P. J., Mesulam, M. M., Parrish, T. B., & Paller, K. A. (2008). Familiarity and conceptual priming engage distinct cortical networks. *Cerebral Cortex, 18,* 1712–1719.

Wagner, A. D., Koutstaal, W., Maril, A., Schacter, D. L., & Buckner, R. L. (2000). Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex, 10,* 1176–1184.

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences, 9,* 445–453.

Weiskrantz, L. (1956). Behavioural changes associated with ablation of the amygdaloid complex in monkeys. *Journal of Comparative Physiology and Psychology, 49,* 381–391.

Weldon, M. S. (1991). Mechanisms underlying priming on perceptual tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 526–541.

Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews, 41,* 88–123.