

Word Imageability Enhances Association-memory by Increasing Hippocampal Engagement

Jeremy B. Caplan and Christopher R. Madan

Abstract

■ The hippocampus is thought to support association-memory, particularly when tested with cued recall. One of the most well-known and studied factors that influences accuracy of verbal association-memory is imageability; participants remember pairs of high-imageability words better than pairs of low-imageability words. High-imageability words are also remembered better in tests of item-memory. However, we previously found that item-memory effects could not explain the enhancement in cued recall, suggesting that imageability enhances association-memory strength. Here we report an fMRI study designed to ask, what is the role of the hippocampus in the memory advantage for associations due to imageability? We tested two alternative hypotheses: (1) Recruitment Hypothesis: High-imageability pairs are remembered better because they recruit the underlying hippo-

campal association-memory function more effectively. Alternatively, (2) Bypassing Hypothesis: Imageability functions by making the association-forming process easier, enhancing memory in a way that bypasses the hippocampus, as has been found, for example, with explicit unitization imagery strategies. Results found, first, hippocampal BOLD signal was greater during study and recall of high- than low-imageability word pairs. Second, the difference in activity between recalled and forgotten pairs showed a main effect, but no significant interaction with imageability, challenging the bypassing hypothesis, but consistent with the predictions derived from the recruitment hypothesis. Our findings suggest that certain stimulus properties, like imageability, may leverage, rather than avoid, the associative function of the hippocampus to support superior association-memory. ■

INTRODUCTION

The hippocampus plays a prominent role in memory for associations—remembering not only which items were studied but also which items were paired with one another. However, there do seem to be factors that make association-memory either less hippocampal-dependent or even completely nonreliant on the hippocampus. For example, after study of a set of pairs, A–B, C–D, E–F, ..., one can test association-memory with cued recall (given A, recall B). Alternatively, one can test association-memory with associative recognition: given two studied items, judge whether the items had been presented together (e.g., C–D) or not (e.g., A–F). Although hippocampal activity can be seen during successful study of word pairs later tested with associative recognition (Jackson & Schacter, 2003), this activity may not be necessary; associative recognition of pairs wherein both items (A and B) are of the same material (e.g., both words or both faces) may remain intact after hippocampal damage, despite impaired associative recognition of mixed-material pairs such as nouns paired with faces (Mayes, Montaldi, & Migo, 2007). Other findings suggest that if a pair can be processed like a single item, so-called “unitization” (Murray & Kensinger, 2013),

the association-memory functions can shift outside the hippocampus (Bader, Opitz, Reith, & Mecklinger, 2014; Diana, Yonelinas, & Ranganath, 2008; Quamme, Yonelinas, & Norman, 2007), even for between-domain associations like color–object pairs (Staresina & Davachi, 2010). On the other hand, recall, even of items, is impaired in hippocampal amnesics (Mayes et al., 2007). Thus, unlike associative recognition, cued recall, even of same-material associations, may be hippocampal-dependent.

One of the most well-known factors influencing association-memory is imageability, a normed measure of respondents’ judgments about their ability to form a mental image of a given word, which is also highly correlated with another word property, concreteness (Paivio, Yuille, & Madigan, 1968). Pairs composed of high-imageability words are remembered better than pairs composed of low-imageability words (Paivio, 1969, 1971, 1986). The reason imageability can increase memory in cued-recall tasks was ambiguous. On the one hand, imageability of individual words could strengthen the association between word A and word B of a pair, leading to better cued recall. On the other hand, cued recall could also be enhanced by the (high) imageability of each individual word, without acting on the association. For example, compared with low-imageability words, a high-imageability word A could act as a better probe to cue any type of memory, or a high-imageability word B could be easier to produce as a

University of Alberta

Both authors contributed equally to this work.

memory target. In both cases, cued recall would be better for high-imageability pairs, but in the first case, due to association-memory processes, whereas in the other cases, due to memory for the individual items. Madan, Glaholt, and Caplan (2010) experimentally disentangled these two effects, using a model-based approach. In their design, participants were asked to learn both “pure” pairs, consisting of two high-imageability or two low-imageability words, and “mixed” pairs, consisting of one high- and one low-imageability word. Pairs were tested equally often in the forward (given A, recall B) and backward (given B, recall A) directions. Then they fit a probabilistic, mathematical model to the data. The model included free parameters that represented the effects of imageability on cued-recall accuracy: namely, whether imageability acts on the strength of the association or on item-memory, regardless of whether the item was paired with a high- or low-imageability word and how the pair is tested (forward or backward). They found that imageability (controlling for word frequency) did indeed enhance cued recall by acting on association strength, with no significant effect on target retrievability.

Imageability effects are already proving difficult to explain in the context of other tasks such as lexical decision and item-memory paradigms, in both neuroimaging and behavioral studies (elaborated in the Discussion). Rather than explain cognitively why imageability enhances association-memory, we take a different approach. We test the specific role of the hippocampus in this effect, if any. As described above, many other studies have found that task conditions that are easier or more accurate appear to be so because those better-performed conditions are independent of the hippocampus (e.g., unitization strategies or intrainem associations like compound words). Thus, we wondered whether the effect of imageability on association-memory is another such example, where high-imageability materials shift the memory demand outside the hippocampus. Alternatively, cued recall of unrelated word pairs may simply always be hippocampal-dependent (Caplan, Boulton, & Gagné, 2014). If so, then high-imageability stimuli may be more effective in recruiting the hippocampus, enhancing memory by increasing the already-present function of the hippocampus, as suggested by Peters and Daum (2008). In line with this, Bonner, Price, Peelle, and Grossman (2016) provided evidence that the hippocampus may be partly responsible for imageability effects even in nonmemory tasks. To our knowledge, support for memory enhancement acting via further hippocampal engagement would be a novel kind of finding concerning the role of the hippocampus in memory behavior.

We recorded brain activity (BOLD-fMRI) while participants studied pairs composed of high-imageability words (HH pairs) and low-imageability words (LL pairs), using the materials from Madan et al. (2010), and were tested with a covert cued-recall procedure, detailed in the Methods section. We test two alternative hypotheses:

- (1) Recruitment Hypothesis: The hippocampus is more engaged during study of high-imageability pairs, leading to increased cued-recall accuracy (Figure 1) and
- (2) Bypassing Hypothesis: The memory boost for high-imageability pairs is due to processes that do not rely on the hippocampus, but rather, supplement its function (Figure 2).

Simulation-generated Predictions

Rather than rely on intuition to generate predictions from the two hypotheses, we ran numerical, Monte Carlo simulations, implementing the Recruitment Hypothesis (Figure 1) and the Bypassing Hypothesis (Figure 2), respectively. In both simulations, we assumed that hippocampal activity could be high or low: model hippocampal activity values were drawn from a mixture of two Gaussian distributions, $N(\mu, \sigma)$, that differed only in their mean activity level, denoted $\eta_+ = N(1.5, 0.1)$ and $\eta_- = N(0.5, 0.1)$, respectively. The plots were generated by simulating activity for 200,000 pairs.

The essence of the Recruitment Hypothesis is that hippocampal activity exclusively drives memory outcome: If

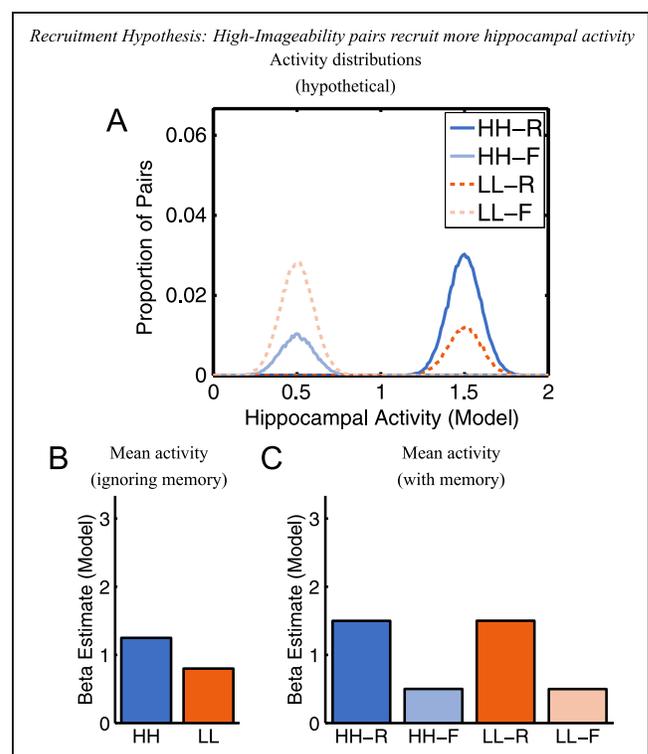


Figure 1. Numerical simulation of the Recruitment Hypothesis with respect to hippocampal activity. (A) Hypothetical distributions of hippocampal activity (proportion of pairs as functions of hippocampal activity with respect to baseline) values as a function of pair imageability and memory outcome. (B) Measured mean activity levels as a function of pair imageability, if memory outcome is ignored (computed from the values from A). (C) Measured mean activity levels as a function of both pair imageability and memory outcome (also computed from the values from A). For more details, see main text and compare with Figure 2 and results, Figures 5 and 6.

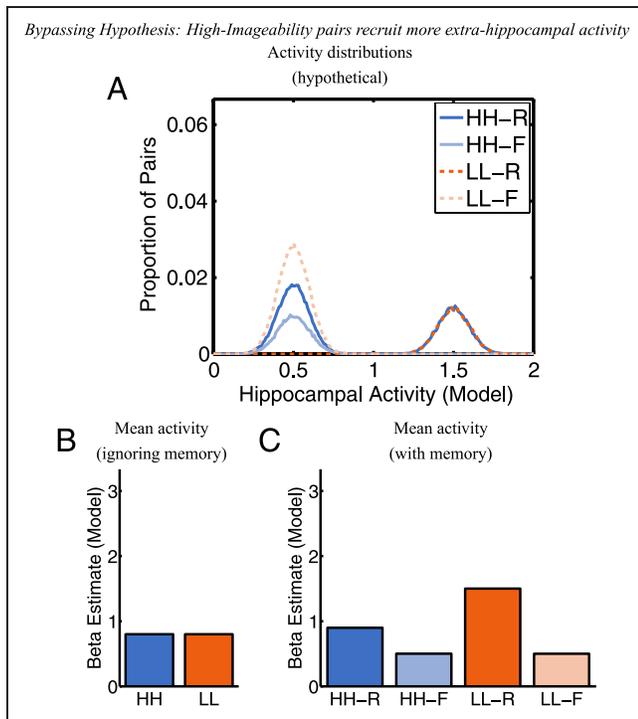


Figure 2. Numerical simulation of the Bypassing Hypothesis with respect to hippocampal activity. (A) Hypothetical distributions of hippocampal activity (proportion of pairs as functions of hippocampal activity with respect to baseline) values as a function of pair imageability and memory outcome. (B) Measured mean activity levels as a function of pair imageability, if memory outcome is ignored (computed from the values from A). (C) Measured mean activity levels as a function of both pair imageability and memory outcome (also computed from the values from A). Note that the high-activity portions of the distributions for HH-R and LL-R overlap and thus nearly occlude one another in A. For more details, see main text and compare with Figure 1 and results, Figures 5 and 6.

a pair is studied with high hippocampal activity (drawn from η_+), it will be remembered, and if studied with low hippocampal activity (drawn from η_-), it will be forgotten. In this view, imageability can act only by invoking hippocampal activity; namely, by increasing the likelihood that a (high-imageability) pair was studied with hippocampal activity drawn from η_+ . Thus, we assumed that hippocampal activity for each pair was drawn from η_+ or η_- , but that high pairs were more likely to be drawn from η_+ , as plotted in the probability distribution function (Figure 1A), when accuracy was fixed to 0.75 and 0.30 for HH and LL pairs, respectively. When one computes average activity during study of HH and LL pairs regardless of memory outcome (Figure 1B), mean hippocampal activity is greater for HH than LL pairs. What is interesting is that when one controls for memory outcome, computing average hippocampal activity separately for later-recalled and later-forgotten pairs (Figure 1C), the effect of pair type goes away, and all that is left is an effect of memory outcome. Thus, Figure 1 (B, C) depicts the pattern of hippocampal activity expected if the Recruitment Hypothesis

holds, and imageability effects are mediated by hippocampal activity.

The idea behind the Bypassing Hypothesis is that the boost in memory success for high-imageability pairs is not due to hippocampal activity but due to some other, extrahippocampal process that enables some pairs to be remembered. Thus, in this view, a pair can be remembered either because hippocampal activity was high (drawn from η_+) or due to brain activity outside the hippocampus, whether hippocampal activity was high (drawn from η_+) or low (drawn from η_-). To simulate the Bypassing Hypothesis, we therefore assumed that the number of pairs with high hippocampal activity were the same for HH and LL pairs, but that a subset of high-imageability pairs that had strengths drawn from η_- were remembered nonetheless, resulting in higher accuracy. The corresponding probability distribution functions are plotted in Figure 2A. The prediction for hippocampal activity, before memory outcome is taken into account, is no difference in activity between HH and LL pairs (Figure 2B). Critically, the Bypassing Hypothesis predicts that the hippocampal subsequent memory effect should be smaller for high- than for low-imageability pairs, predicting an Imageability \times Memory interaction (Figure 2C). Thus, Figure 2 (B, C) depicts the pattern of hippocampal activity expected if the Bypassing Hypothesis holds. Observe that the Bypassing Hypothesis predicts even greater hippocampal activity during study of LL-remembered than HH-remembered pairs, because the HH pairs can be remembered even when hippocampal activity was low. Note that by assumption, the bypassing account implies no effect of imageability on hippocampal activity. A variant of the Bypassing Hypothesis might be that the extrahippocampal support, in reducing the need for the hippocampus to contribute to memory, even enables the hippocampus to reduce its activity during study of HH pairs. This would still produce an interaction when memory is incorporated into the analysis but would lead one to predict more hippocampal activity during study of LL than HH pairs, even more different than the corresponding prediction based on the Recruitment Hypothesis (Figure 1B).

An alternative formulation of the Recruitment Hypothesis is to assume that high-imageability pairs result in a Gaussian distribution that is right-shifted. When we simulated the hypothesis in this way (not shown), the predictions were quite similar. Thus, to select between the Recruitment and Bypassing hypotheses, we analyzed hippocampal activity both without and with memory outcome taken into account and compared those results with the predictions in Figures 1 and 2.

Although we developed our hypotheses for brain activity during the study phase, various lines of evidence have suggested that memory that is hippocampal-dependent during study is also hippocampal-dependent at test (e.g., Sparks, Lehmann, & Sutherland, 2011; Epp et al., 2008; Gilboa, Winocur, Grady, Hevenor, & Moscovitch,

2004; Nadel, Samsonovich, Ryan, & Moscovitch, 2000; Nadel & Moscovitch, 1997). We analyzed activity during the test phase to test whether the hypotheses at study found consistent support at test as well. Although our hypotheses are formulated only in terms of hippocampal activity, we broke the hippocampus down by left versus right side to check for material specificity. One may expect that left hippocampus contributes more to verbal processing and right hippocampus contributes more to visuospatial processing (e.g., Schacter & Wagner, 1999; Milner, 1954), although the evidence for this is mixed (e.g., Saling, 2009; Bell & Davies, 1998), especially in neuroimaging studies of healthy participants (Dalton, Hornberger, & Piguet, 2016). Insofar as our task is verbal, left hippocampus should be important, but if visual imagery plays a role in memory in our task, then the right hippocampus may be critical. In addition, increasing attention has been paid to the anterior–posterior (long) axis of the hippocampus, but without a clear consensus as to what the critical factors are (e.g., Poppenk, Evensmoen, Moscovitch, & Nadel, 2013; Schacter & Wagner, 1999). For example, Dalton et al. (2016) found left anterior hippocampal activity related to encoding of compound words (tested with associative recognition) and right-sided activity that was more posterior related to encoding of nonverbal, visual associations. Thus, if the imageability effect is related to memory processes involved in nonverbal, visual memory, we might expect more posterior than anterior right hippocampal activity. Therefore, in case memory function in our cued-recall task was either more anterior or posterior, we also broke down each hippocampus into an anterior and a posterior subregion. Following the hippocampal ROI analyses, we conducted whole-brain, exploratory analyses to seek which, if any, regions beyond the hippocampus might explain both memory, in general, and the imageability effect.

METHODS

Participants

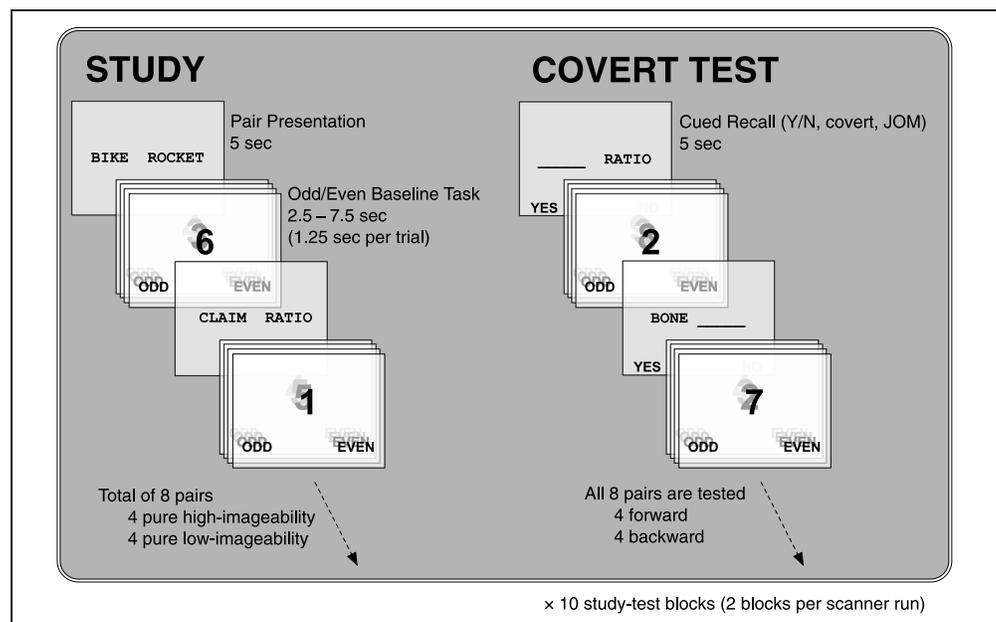
A total of 31 volunteers (age = 23.1 ± 3.4 years; 13 women, 18 men) participated in our study and were paid \$25/hr. All participants were required to have learned English before the age of 6, be right-handed as measured by the Edinburgh Handedness Inventory (Oldfield, 1971), laterality quotient, mean \pm SD = $90.0 \pm 10.9\%$, normal or corrected-to-normal vision, no neurological or psychiatric disorders, and no contraindications to MR scanning. Participants gave written informed consent before beginning the study, which was approved by a University of Alberta research ethics board.

Behavioral Methods

Overall Design

Participants studied and were tested on multiple sets composed of eight word pairs each (Figure 3). To increase the likelihood that the task was hippocampal-dependent, we used a cued-recall memory test. Although vocal recall is possible (Henson, Shallice, Josephs, & Dolan, 2002), it is challenging to perform reliably against the scanner noise, especially in high field strengths such as here. We therefore asked participants, given a cue word, to covertly recall the associated word, similar to other studies (e.g., Horner, Bisby, Bush, Lin, & Burgess, 2015; Bosch, Jehee, Fernández, & Doeller, 2014; Hannula, Libby, Yonelinas, & Ranganath, 2013; Wang, Yonelinas, & Ranganath, 2013; Wolosin, Zeithamova, & Preston, 2012; Meltzer & Constable, 2005). The covert response method was used in each test block during scanning. To verify covert cued-recall success, after the scan, participants were given the complete list of pairs again, with identical pairings and test order and responded overtly, typing on a computer keyboard. Only

Figure 3. Illustration of the experimental procedure. This schematic figure illustrates the tasks performed within the MR scanner. Not depicted are an initial practice phase and a final set of overt cued-recall tests that took place outside the scanner.



the 20 participants with the highest concordance between covert and overt responses (minimum correlation = 0.7) were retained (see below for details). This enabled us to base the data analyses (subsequent memory effects) on the covert recall responses, which produced accuracies farther from floor levels (cf. Hannula et al., 2013; Wolosin et al., 2012).

Materials

Eight-pair lists were constructed pseudorandomly for each participant, from two pools of nouns used in a previous experiment (further stimulus details are reported in Madan et al., 2010): a set of 110 high-imageability nouns and a set of 110 low-imageability nouns. Pools were matched on word frequency (high: 21 ± 14 , low: 20 ± 15 words per million), word length (high: 4.98 ± 0.79 , low: 5.00 ± 0.79 letters), mean positional bigram frequency, and orthographic neighborhood size. Pairwise semantic similarity, measured with latent semantic analysis (Landauer & Dumais, 1997), also did not differ significantly between high- and low-imageability pairs (LSA $\cos(\theta)$, high: 0.10 ± 0.14 , low: 0.12 ± 0.15). Consulting the Brysbaert and New (2009) database, the high- and low-imageability word pools did not differ in contextual diversity either, $t(218) = 1.56, p = .12$.

We use the term “imageability” rather than “concreteness,” although most prior studies, including those cited in the Introduction, have referred to concreteness. Concreteness is quantified in a similar manner as imageability based on subjective ratings, and concreteness and imageability are highly correlated with one another. Still, some words have high concreteness but low imageability ratings, and others have low concreteness but high imageability ratings. Because it was unclear to us precisely how imageability and concreteness differ from one another, when we initially developed the stimulus pools, we restricted our pools to include only words that had consistent imageability and concreteness ratings (Madan et al., 2010). In this sense, in this study, imageability and concreteness are interchangeable; thus, we prefer the term imageability, which we suggest is slightly less subjective than concreteness, but it should be noted that our findings may be equally well understood in terms of concreteness as imageability.

Each pair was composed either of two high-imageability words (HH pairs) or two low-imageability words (LL pairs). Each study set consisted of four HH and four LL pairs; pairs were presented for 5 sec (see Figure 3). All words were presented the center of the screen in a white Courier New font, which ensured fixed letter width, on a black background. The task was designed using the Python Experiment-Programming Library (Geller, Schleifer, Sederberg, Jacobs, & Kahana, 2007).

Procedure

In the scanner, stimuli were back-projected on a screen and viewed through a mirror. Participants made responses

using two button boxes, one in each hand, using the respective index fingers. The mapping of left/right index fingers to odd/even, in the active baseline task, and yes/no, in the memory judgment, was displayed during the task to remind the participants how to respond and counterbalanced across participants.

Participants were presented with 10 experimental blocks of eight pairs each (Figure 3). Each pair was presented centrally for 5 sec. Each of five scanning runs consisted of two blocks (study–test sets), lasting 6.8 min each.

Because the hippocampus is often found to be active during unfilled rest periods, we included an active baseline task, asking participants to judge numbers as odd or even. Interleaved between study pairs and tests probes, this task functioned as an implicit baseline, with presumably minimal hippocampal activity (as shown by Stark & Squire, 2001). In this task, participants had to judge if a presented number, from 1 to 9, was even or odd. Participants were given 1.25 sec per trial. In the practice, responses were made using the “Z” and “/” keys on a computer keyboard. A variable number, ranging from 2 to 6 (pseudorandomly), of baseline trials were presented between pairs as a filled intertrial interval, corresponding to a duration of 2.5–7.5 sec. Following the study phase, 12 more baseline trials were included as an end-of-list distractor to attenuate recency effects.

The study phase was followed by a set of eight covert cued-recall trials. In each test trial, participants were given a single word probe from the just-studied set and asked to recall its associate covertly. The probe word was presented with a blank line either to the right or left of the word (forward and backward testing directions, respectively). Participants were asked to make a metamemory judgment, responding with a button press indicating either “yes,” they could remember the associate of the cue word, or “no” if not, but participants did not overtly recall the target word while in the scanner. This was repeated for all pairs in a pseudorandomized sequence, with participants allotted 5 sec to make their response. Test direction was pseudorandomized such that an equal number of trials were tested in the forward and backward directions. As during study, two to six (pseudorandomly) baseline trials were presented between successive recall trials and an additional 12 baseline trials after all eight pairs had been tested.

After all five scanning runs were completed, the structural image was acquired. After exiting the scanner, to validate the covert cued recall test, participants were tested on all 80 pairs with overt cued recall, in the same overall order and with same probe words as in the main experiment. Participants typed the associated word on the computer keyboard, and it appeared on the blank line as they typed. After each response, a 150-msec blank screen was presented before the next cued-recall trial appeared. There was no maximum RT, but participants were asked to type “PASS” if they could not recall the associated word. Words had to be spelled correctly to be scored as correct.

Before entering the scanner, in a practice phase, participants were familiarized with all aspects of the task, with identical timing as the scanned tasks and overt recall of those eight practice pairs. Practice words were not reused in the experiment.

Neuroimaging Methods

Image Acquisition

Images were acquired on a Varian (Palo Alto, CA) Unity Inova 4.7 T scanner using a single-channel transverse electromagnetic head coil (MR Instruments, Inc., Minnetonka, MN). Functional images were collected using single-shot gradient-echo EPI sequence, with repetition time = 2500 msec, echo time = 28 msec, flip angle = 75°, field of view = 240 × 240 mm², 39 slices with 3 × 3 × 3 mm³ (isotropic) voxels. Slices were oriented in axial-oblique plane, aligned parallel to the AC–PC. Anatomical images were collected using a MPRAGE sequence, repetition time = 1506 msec, echo time = 3.71 msec, field of view = 240 × 192 × 192 mm³ (1 × 1 × 1.5 mm³ voxels), with slice orientation matching the functional data.

fMRI Preprocessing

The first five volumes of each scanning run were discarded to allow for stabilization of the BOLD signal. The image time series was slice time-corrected, realigned, and corrected for the interaction of motion and distortion (unwarped). The T1 images were normalized into standard anatomical space using DARTEL. The resulting deformation field was applied to the EPI images, which were then smoothed with a Gaussian kernel of 8-mm FWHM using SPM8 (www.fil.ion.ucl.ac.uk/spm).

fMRI Data Analysis

First-level models were constructed for each participant. We conducted these in two different ways to address two different kinds of questions, as laid out in the Introduction. First, we ran the first-level analyses separating pairs only according to imageability (high vs. low) to ask the first kind of question: Is the hippocampus (hippocampal ROI analyses) or any other region (whole-brain analyses) more active during study of high- than low-imageability pairs, without concern for memory outcome (cf. Figures 1B and 2B)? These two conditions, for both study and test, were first modeled as separate regressors, using the SPM canonical hemodynamic response function. We then ran a second set of analyses also separating pairs according to memory performance, based on covert cued recall; namely, a subsequent memory effect during study or retrieval success effect during test (Figures 1C and 2C). The resulting four conditions (high-remembered, high-forgotten, low-remembered, low-forgotten) for each of study and test were modeled as separate regressors. As explained in the

Introduction and Figures 1 and 2, both analyses are necessary to get the complete picture, since, for example, the Recruitment Hypothesis predicts main effect in the two-condition analysis but not in the four-condition analysis.

In second-level analyses, activity related to imageability, regardless of memory success, was identified by contrasting high and low pairs (main effect of imageability). Successful association formation, regardless of pair imageability, was identified by contrasting remembered and forgotten (main effect of memory). To identify brain regions that showed differential participation in successful association formation (study) or retrieval (test) as a function of imageability, we contrasted activity associated with the subsequent memory effect or retrieval success effect in high versus low pairs by applying both respective interaction contrasts (i.e., [High-Remembered – High-Forgotten] > [Low-Remembered – Low-Forgotten] and the opposite contrast).

Hippocampal ROI Analyses

As our primary research question concerned the role of the hippocampus in association-memory and potential influence of imageability, we performed a series of planned comparisons using anatomically defined hippocampal ROI analyses. Anticipating lateralization effects and anterior–posterior differences (e.g., Poppenk et al., 2013), we defined four hippocampal ROIs, treated as a 2 × 2 (Hemisphere × Long Axis) factorial design in the analyses. Using the maximum probability atlas developed by Hammers et al. (2003), we extracted the bilateral hippocampal region based on MNI coordinates. Hammers et al. (2003) delineated the hippocampus using several boundaries, including the uncal sulcus, sulcus hippocampalis, amygdala, and parahippocampal gyrus (for further details, see appendix of Hammers et al., 2003). The caudal boundary of the anterior section was defined as $y = -21$ mm, corresponding to the position of the uncal apex, the anterior–posterior boundary in the landmark-based segmentation method described in Poppenk et al. (2013). Thus, our ROI analyses are based on four hippocampal regions, treated as a 2 × 2 (Hemisphere × Long Axis) factorial design. For the ROI analyses, we extracted the beta values from the respective first level for all voxels within each ROI for each condition. We then averaged across all voxels for each ROI and condition, which were then used as the input to the ANOVA analyses. The four ROIs ranged from 999 to 2079 mm³ (number of voxels were left anterior: 57, left posterior: 39, right anterior: 77, right posterior: 37), with the anterior ROIs being larger in volume than the posterior ROIs (Figure 5).

Whole-brain Analyses

To investigate extrahippocampal contributions to memory and the memory advantage for high-imageability pairs, we also performed exploratory, whole-brain analyses. AFNI's

3dClustSim (Cox, 1996) was used to estimate the overall probability of false positives within the 3-D whole-brain search volume to achieve an α of .05 corrected for multiple comparisons ($p < .001$ and $k \geq 19$ voxels). If a cluster was judged to span more than one anatomical region, a watershed algorithm was applied to parcel the cluster into contiguous subclusters (upenncmroi.wpengine.com/shed/), as described in Satterthwaite et al. (2013). This algorithm first identifies local maxima within the cluster and then computes the extent of the subclusters using an expansion process modeled after how water would drain into an inverted topology of the activation map. These resulting subclusters were required to have a minimum size of 19 voxels each. Localizations of significant clusters were determined by reference of the structural MRIs to the Duvernoy (1991) atlas.

RESULTS

Behavioral Results

All participants performed well on the baseline task, $M \pm 95\%$ CI = $90.4 \pm 1.8\%$. We observed a large advantage for high-imageability pairs in the overt, postscan cued-recall test, $t(30) = 10.13, p < .001, M_{\text{high}} = 46.0 \pm 2.8\%, M_{\text{low}} = 16.8 \pm 2.8\%$, as expected. This imageability effect was also reflected in the covert (in-scanner) responses, covert: $t(30) = 8.74, p < .001, M_{\text{high}} = 69.3 \pm 2.7\%, M_{\text{low}} = 44.9 \pm 2.7\%$; thus, the covert responses also showed an enhancing effect of imageability. As expected, covert recall accuracy exceeded overt recall accuracy, because the covert test immediately followed the study phase (and distractor), whereas the overt test was delayed (end of session) and occurred after all of the study sets had been studied and tested.

We restricted our main analyses to participants for whom the covert recall performance could be validated by the postscan overt recall performance (but an inclusive version of the hippocampal ROI analyses produced the same central results, as noted in their respective Results sections). We calculated Yule's Q to characterize the relationship between covert and overt recall accuracy; Q is a special case of the gamma correlation which is appropriate for dichotomous values (correct/incorrect) and, like Pearson correlation, controls for differences in mean accuracy (marginals). Note that means, confidence intervals, and statistical tests are computed on log-odds transformed values to satisfy the assumption of normality. Covert recall accuracy was highly correlated with overt cued-recall accuracy (high: 95% CI of Yule's $Q = [0.76, 0.90]$; low: $Q = [0.72, 0.88]$) and did not differ by imageability, $t(30) = 0.98, p = .33$. This suggests that the covert responses are a reasonably good stand-in for cued-recall accuracy. We first removed participants who had no correct overt recalls to either high or low pairs, which makes it impossible to conduct subsequent memory effect analyses, excluding four participants. Of the

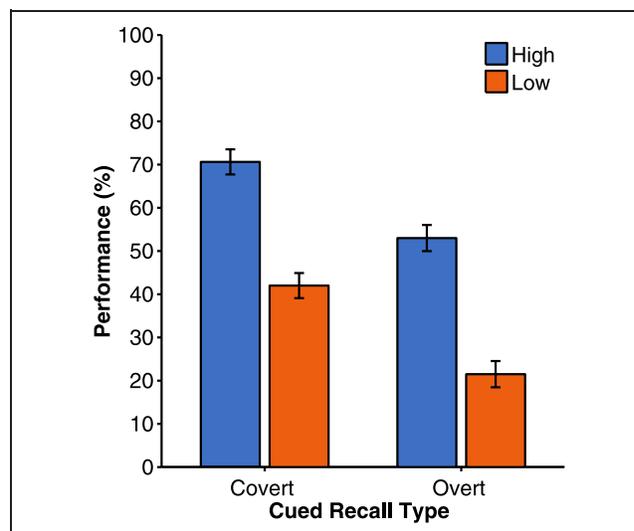


Figure 4. Cued-recall accuracy for the final sample ($n = 20$), as a function of pair imageability (high- vs. low-imageability) and test (covert vs. overt cued recall). The covert test immediately followed the study phase and end-of-list distractor (baseline task), whereas the overt test was administered at the end of the testing session, outside the scanner, accounting for the relatively lower accuracy in the overt test. Error bars are 95% confidence intervals based on *SEM*, after correcting for subject variability (Loftus & Masson, 1994).

remaining participants, we excluded participants for whom either $Q_{\text{high}} \leq 0.70$ or $Q_{\text{low}} \leq 0.70$ (seven participants), leaving $n = 20$ participants in the final sample. This subsample still demonstrated strong imageability effects in both covert [$t(19) = 9.63, p < .001, M_{\text{high}} = 70.6 \pm 2.9\%, M_{\text{low}} = 42.0 \pm 2.9\%$] and overt [$t(19) = 10.17, p < .001, M_{\text{high}} = 53.0 \pm 3.0\%, M_{\text{low}} = 21.5 \pm 3.0\%$] cued-recall tests (Figure 4). As expected, the covert–overt correlations increased (95% CI for high: $Q = [0.88, 0.93]$ and low: $Q = [0.89, 0.93]$) and still did not differ significantly by imageability, $t(19) = 0.27, p = .79$.

Hippocampal ROI Results

We analyzed brain activity within the four hippocampal ROIs (Hemisphere \times Long Axis). Described in greater detail in the Methods, beta values from the first-level analyses were averaged together for voxels within each anatomically delineated ROI and then entered into ANOVAs and correlation analyses. The ROIs themselves and the corresponding beta-value plots are illustrated for the Study and Test phases in Figures 5 and 6, respectively.

Engagement of the Hippocampus without Regard to Memory Outcome

Do high-imageability stimuli engage the hippocampus more than low-imageability stimuli? This is the first condition of the Recruitment Hypothesis: That high-imageability pairs are remembered better due to greater engagement of the hippocampus (Figure 1B vs. Figure 2B). We therefore

start by analyzing activity during the study and test phases without taking into account memory outcome (left-hand bar graphs in Figures 5 and 6, respectively), conducting an ANOVA with a $2 \times 2 \times 2 \times 2$ design: Hemisphere (left, right) \times Long Axis (anterior, posterior) \times Phase (study, test) \times Imageability (high, low). The main effect of Imageability was significant, $F(1, 19) = 11.28, MSE = 1.64, p < .01$, with greater hippocampal activity during study and test of high- than low-imageability pairs. This confirms the first prediction of the Recruitment Hypothesis and is inconsistent with the Bypassing Hypothesis.

This effect varied to some degree with hippocampal sub-region; the main effect of Imageability was qualified by a significant Imageability \times Hemisphere interaction, $F(1,$

$19) = 19.94, MSE = 0.19, p < .01$. Simple effects (separated by hemisphere) found that the main effect of Imageability was significant in both hemispheres but slightly more robust in the left hippocampus ($M_{\text{high}} = 1.96, M_{\text{low}} = 0.92$, main effect of Imageability: $F(1, 19) = 8.00, MSE = 5.43, p < .05$) than in the right hippocampus ($M_{\text{high}} = 1.08, M_{\text{low}} = 0.44$, main effect of Imageability: $F(1, 19) = 6.35, MSE = 2.25, p < .05$).

The only other significant effects did not include Imageability and thus do not speak directly to the hypotheses. For completeness, these were the main effect of Phase, $F(1, 19) = 156.87, MSE = 9.85, p < .01$, with greater activity during Study than Test, and the interactions Phase \times Hemisphere, $F(1, 19) = 15.5, MSE = 0.57, p < .01$,

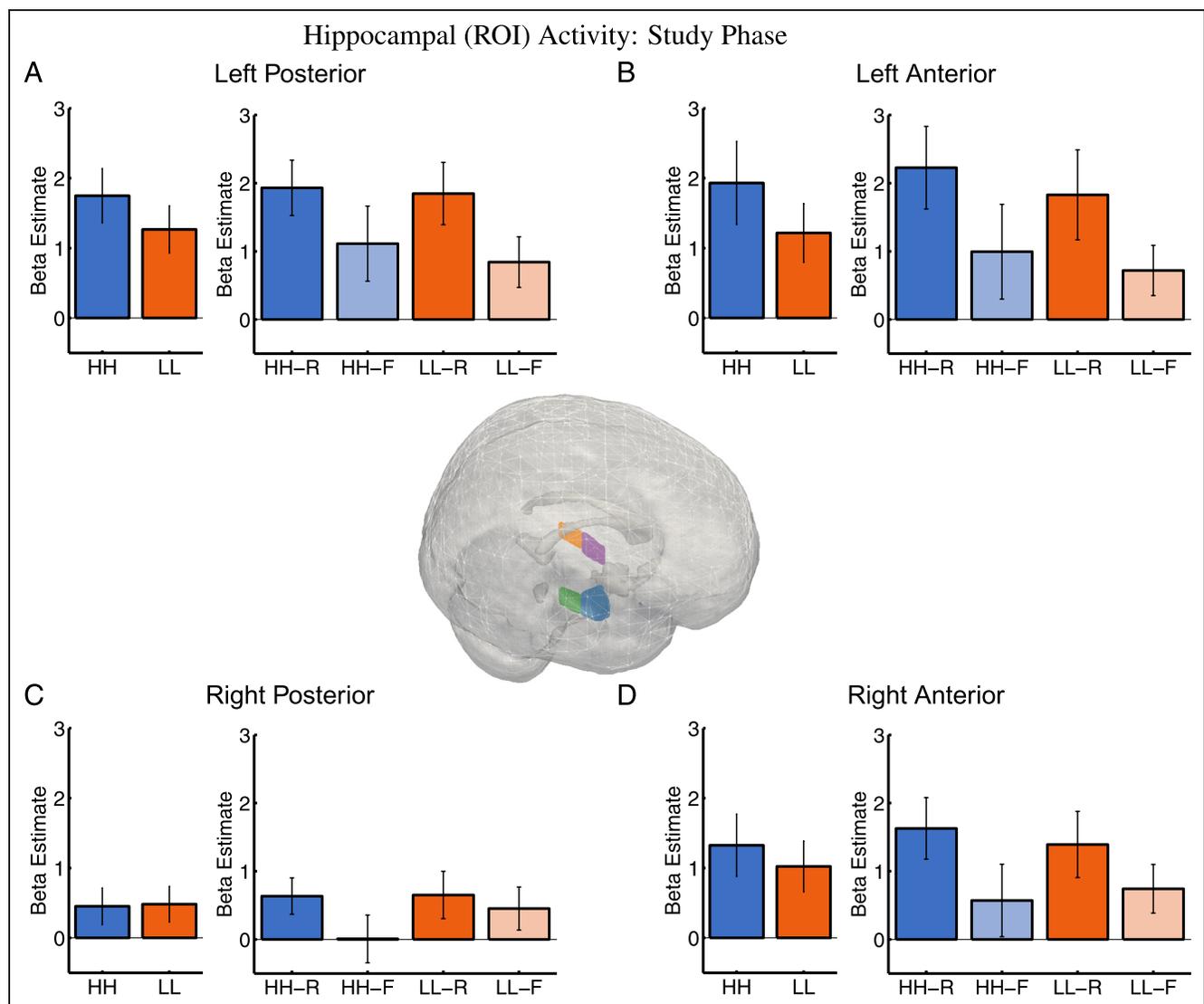


Figure 5. Hippocampal ROI analysis during the Study phase. Middle, 3-D rendering of the four hippocampal ROIs within the brain, as extracted from Hammers et al. (2003) and segmented along the long-axis based on Poppenk et al. (2013). The rendering is viewed from a right, superior perspective and made using ParaView 4.2.0 (Kitware Inc., Clifton Park, NY), following the procedure described in Madan (2015). For each ROI, the mean (and *SE*) beta estimates are plotted first, during study of high-imageability (HH) and low-imageability (LL) pairs, derived from the first analysis, without taking memory outcome into account, and then, during study of high- and low-imageability pairs broken down by subsequent (covert response) memory outcome (R = later remembered; F = later forgotten), derived from the second analysis that incorporated memory: (A) left posterior hippocampal ROI, (B) left anterior hippocampal ROI, (C) right posterior hippocampal ROI, (D) right anterior hippocampal ROI.

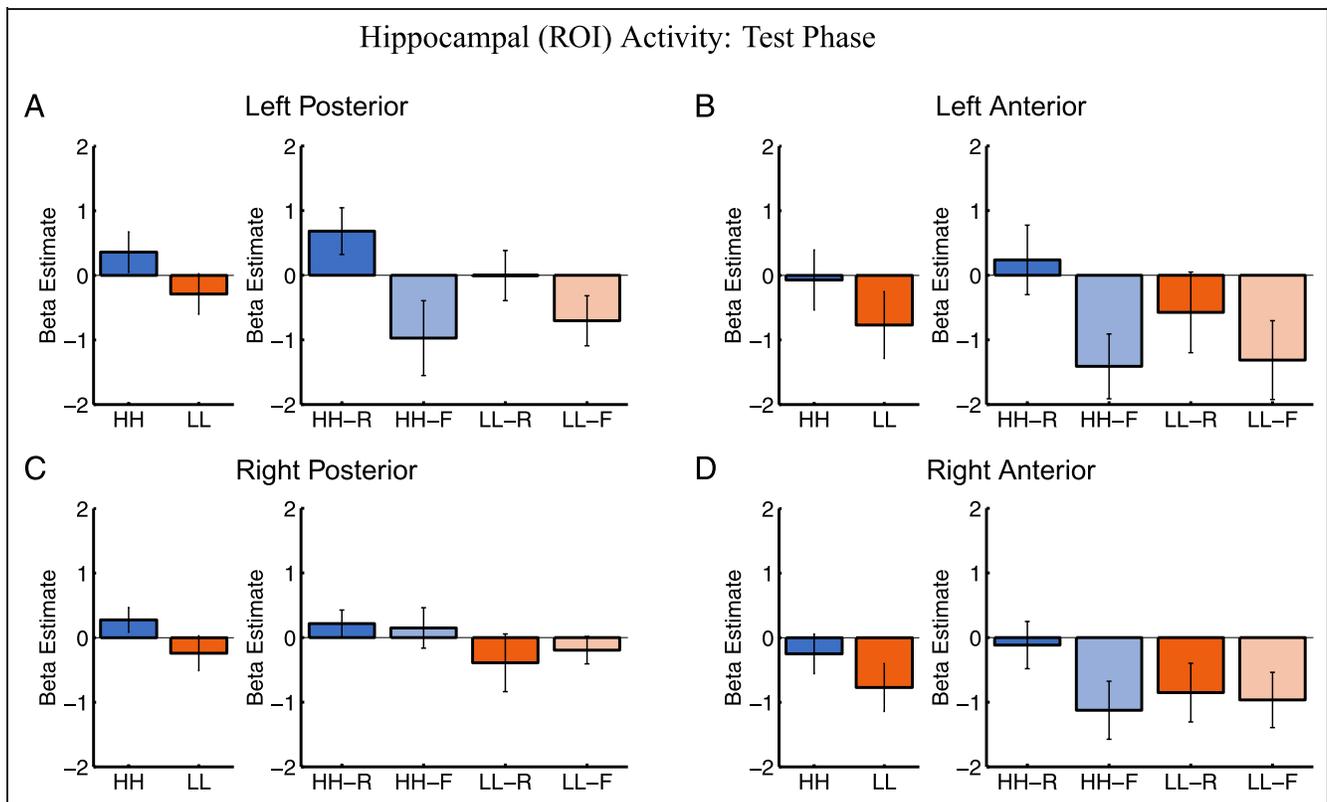


Figure 6. Hippocampal ROI analysis during the test phase, ROIs selected as in Figure 5. For each ROI, the mean (and *SE*) beta estimates are plotted first, during test of high-imageability (HH) and low-imageability (LL) pairs, derived from the first analysis, without taking memory outcome into account, and then, during test of high- and low-imageability pairs broken down by (covert response) memory outcome (R = remembered; F = forgotten), derived from the second analysis that incorporated memory: (A) left posterior hippocampal ROI, (B) left anterior hippocampal ROI, (C) right posterior hippocampal ROI, (D) right anterior hippocampal ROI.

Phase \times Long Axis, $F(1, 19) = 5.84$, $MSE = 2.64$, $p < .05$, and the three-way interaction Phase \times Hemisphere \times Long Axis, $F(1, 19) = 11.00$, $MSE = 0.23$, $p < .01$. Finally, two trend effects should be mentioned for completeness: the main effect of Hemisphere, $F(1, 19) = 3.04$, $MSE = 3.96$, $p = .098$, and the interaction Phase \times Hemisphere \times Imageability, $F(1, 19) = 3.21$, $MSE = 0.15$, $p = .089$, approached significance.¹

Thus far, the results are consistent with the Recruitment Hypothesis (Figure 1B), especially on the left side, both during study (Figure 5) and during test (Figure 6).

Influence of hippocampal activity on memory outcome. Greater hippocampal activity during study and cued recall of high- than low-imageability pairs could occur even if increased activity were not responsible for the high-imageability advantage in memory or even if the task were not hippocampal-dependent. In other words, hippocampal activity, including the difference between high- and low-imageability pairs, could be superfluous activity, irrelevant to memory. Therefore, we next incorporated memory outcome, namely, accuracy on the covert cued recall test (right-hand bar graphs in Figures 5 and 6). As described in the Methods, this was based on a different GLM first-level model than the analysis in the previous section. We conducted a $2 \times 2 \times 2 \times 2 \times 2$ ANOVA with design Hemisphere (left, right) \times Long Axis (anterior,

posterior) \times Phase (study, test) \times Imageability (high, low) \times Memory (recalled, forgotten). If the task is indeed hippocampal-dependent as intended, we expect effects involving the Memory factor. If the increase in hippocampal activity drives the advantage for high-imageability pairs, this leads to a very specific and not immediately obvious prediction, as explained in the Introduction; with memory outcome incorporated into the analyses, we expect interactions involving both Imageability and Memory to be non-significant. That is, if hippocampal activity level determines memory outcome and imageability, in turn, influences that hippocampal activity level, then when controlling for subsequent memory outcome, the effects of Imageability should be explained away (Figure 1). If, instead, the Bypassing Hypothesis holds, we expect a significant interaction (Figure 2).

First, there was a significant main effect of Memory, $F(1, 19) = 24.4$, $MSE = 3.95$, $p < .001$, consistent with our expectation that hippocampal activity enhances memory in this task. Greater activity was associated with recalled than forgotten pairs. This effect varied by region, as the Long Axis \times Memory, $F(1, 19) = 5.17$, $MSE = 0.87$, $p < .05$, Hemisphere \times Memory, $F(1, 19) = 20.8$, $MSE = 0.87$, $p < .001$, and Hemisphere \times Long Axis \times Memory, $F(1, 19) = 5.36$, $MSE = 0.29$, $p < .05$, were significant.

Pertinent to the hypotheses, Imageability was not involved in any significant effects, in line with the Recruitment Hypothesis (Figure 1C), but inconsistent with the Bypassing Hypothesis (Figure 2C). The only interaction involving Imageability that even approached significance was Phase \times Hemisphere \times Imageability, $F(1, 19) = 3.89$, $MSE = 0.53$, $p = .063$. Most pertinently, all interactions involving both Imageability and Memory were far from significant (p s $> .18$).²

For completeness, but not speaking directly to our hypotheses, other significant effects were Phase, $F(1, 19) = 11.30$, $MSE = 34.36$, $p < .01$, Phase \times Hemisphere, $F(1, 19) = 4.66$, $MSE = 5.17$, $p < .05$, and Phase \times Hemisphere \times Long Axis, $F(1, 19) = 8.93$, $MSE = 0.78$, $p < .01$, in addition to a nonsignificant trend for Phase \times Long Axis, $F(1, 19) = 4.09$, $MSE = 8.65$, $p = .057$.

Study and Test Phases Analyzed Separately

Finally, we checked whether the main findings would stand when each phase was analyzed on its own. First, in the analysis disregarding memory outcome, study activity showed similar basic effects: Although the main effect of Imageability was a nonsignificant trend, $F(1, 19) = 4.30$, $MSE = 1.25$, $p = .052$, the interaction Imageability \times Hemisphere was still significant, $F(1, 19) = 13.8$, $MSE = 0.16$, $p < .01$, with a significant effect of Imageability in the left but not right hemisphere. Test activity analyzed alone still produced a significant main effect of Imageability, $F(1, 19) = 6.77$, $MSE = 2.10$, $p = .018$, but was no longer qualified by a significant Hemisphere \times Imageability interaction. Thus, the prediction of the Recruitment Hypothesis, of more hippocampal activity for high- than low-imageability pairs, could be confirmed for study and test separately.

Second, in the analysis incorporating memory outcome, Memory was still a significant main effect during the study phase, $F(1, 19) = 7.84$, $MSE = 7.15$, $p < .05$, but the interactions with region no longer reached significance. Importantly, all terms involving Imageability \times Memory were still nonsignificant. For test activity, Memory was also still a significant main effect, $F(1, 19) = 6.52$, $MSE = 6.31$, $p < .05$, and Hemisphere \times Memory was significant, $F(1, 19) = 13.08$, $MSE = 1.34$, $p < .01$, but other interactions with region did not reach significance. Importantly, all terms involving Imageability \times Memory were nonsignificant.

Thus, even when study and test phases were analyzed separately, the main results were still statistically robust.

Bayes Factor Tests of Null Effects

To determine how the nonsignificant interaction effects should be interpreted, we used JASP (JASP Team, 2016) to evaluate the 2×2 Bayesian repeated-measures ANOVA for each phase (study and test) and ROI individually, always assuming uniform prior probabilities. The full model (main effect of Memory and Imageability plus

Memory \times Imageability) had a Bayes factor less than 3 against the null model in all cases; thus, there was never strong evidence for an interaction. The full model had a Bayes factor under 0.3 against the null model for Study–Left–Posterior, Test–Right–Anterior, and Test–Right–Posterior, indicating strong evidence against the interaction at those ROIs. For the remaining ROIs, we considered whether any simpler model (not containing the interaction) would be favored over the full model (with the interaction). At all but one ROI, the model with a main effect of Memory alone had the largest Bayes factor, which was more than three times that for the full model (and greater than 3, itself, against the null model). For the remaining ROI, Test–Right–Posterior, the model with Imageability alone had the largest Bayes factor, which was more than three times that for the full model (but less than 1, itself, against the null model). These results suggest that the lack of significant interaction at all hippocampal ROIs should be viewed as evidence against the interaction rather than a lack of statistical power.

Summary of ROI-based Analyses

In short, the pattern expected based on the Recruitment Hypothesis, a main effect of Imageability without incorporating memory outcome (Figure 1B) and no Imageability \times Memory interaction in the analysis incorporating memory outcome (Figure 1C), was consistent with our results (Figures 5 and 6). In contrast, the pattern expected based on the Bypassing Hypothesis, no main effect of Imageability (or more activity for low- than high-imageability pairs) without incorporating memory outcome (Figure 2B), and an Imageability \times Memory interaction in the analysis incorporating memory outcome (Figure 2C) was inconsistent with our results.

Whole-brain fMRI Results

Next, we examined the broader network of regions supporting memory in this task. We performed exploratory, whole-brain analyses, paralleling the hippocampal ROI-based analyses, except that the multiple-comparison problem was addressed by using a conservative single-voxel significance threshold ($p < .001$) and a cluster-extent threshold (19 voxels; see Methods). Note that we conducted analyses of both the study and test phase activity, but the test phase results were less statistically robust. The analysis involving only the single factor of Imageability revealed more activity during retrieval of high- than low-imageability pairs in medial pFC and precuneus. The analyses involving Imageability and Memory Outcome as factors did not produce effects involving imageability (apart from some apparent activity in white matter). Thus, we confine our report to the study phase analyses. Of greatest interest was whether we could identify any extra-hippocampal regions that might explain the boost in

memory due to imageability, which might indicate some contribution from a bypassing-like process.

Activity without Regard to Memory Outcome

First, in the single, two-level factor, Imageability (high, low) analysis, low-imageability pairs were associated with more activity in left and right superior temporal regions as well as right superior parietal cortex (Table 1). No clusters met our conservative single-voxel and cluster-extent criteria in the direction high imageability > low imageability. Although this may be due to the two conditions being relatively similar to one another, we were also concerned about Type II error and, in particular, the possi-

bility that we might miss activity that was relevant to our imageability manipulation. Therefore, we next relaxed the significance criteria somewhat ($p < .01$, $k \geq 5$) but note that the resulting findings (Table 1) should be interpreted with caution, as they are more liberal than standard recommendations (Woo, Krishnan, & Wager, 2014). Consistent with the hippocampal ROI analyses, hippocampal regions were only found significant in the High > Low contrast and again only in the left hippocampus, consistent with stronger effects in left than right hippocampus in the ROI analyses. It does not appear to be the case that the entire brain was simply more active during study of high- than low-imageability pairs; with these particular thresholds, there was a far greater volume of

Table 1. Whole-brain Exploratory Analysis of the Study Phase before Incorporating Memory Outcome

Cluster No.	MNI (mm)			Peak Voxel Z Score	Volume (No. of Voxels [k])	Region
	x	y	z			
<i>Contrast: High-imageability > Low-imageability</i>						
Strict criteria ($p < .001$, cluster size ≥ 19 voxels):						
[No clusters found]						
Lenient criteria ($p < .01$, cluster size ≥ 5 voxels):						
1	-6	27	6	3.70	14	L corpus callosum (genu)
2	0	39	0	3.23	13	Anterior cingulate cortex
3	-36	-78	39	3.21	43	L angular gyrus
4	-21	-9	-24	3.15	14	L hippocampus (anterior)
5	-24	-24	-9	3.13	10	L hippocampus (posterior)
<i>Contrast: Low-imageability > High-imageability</i>						
Strict criteria ($p < .001$, cluster size ≥ 19 voxels):						
6	-51	3	-21	4.94	55	L superior temporal gyrus
7	15	-63	45	3.88	68	R superior parietal gyrus
8	54	-3	-15	3.79	41	R superior temporal gyrus
Lenient criteria ($p < .01$, cluster size ≥ 5 voxels):						
9	30	-69	-15	3.60	11	R lingual gyrus
10	45	-18	6	3.45	33	R posterior insula
11	45	-27	24	3.44	22	R supramarginal gyrus
12	-39	-69	3	3.15	15	L lateral occipital cortex
13	0	-27	30	3.12	51	Posterior cingulate cortex
14	48	-39	0	3.08	46	R middle temporal gyrus
15	21	48	18	3.07	5	R medial prefrontal cortex
16	-12	-75	36	3.06	13	L precuneus
17	42	6	30	2.90	14	R inferior frontal sulcus

Because few clusters met our strict criteria, $p < .001$ (single-voxel) and cluster size ≥ 19 , this table also reports the more liberal set of results, with single-voxel threshold $p < .01$ and cluster size ≥ 5 voxels (but clusters meeting both criteria are listed only once, and only the strict-criterion cluster sizes are listed in those cases). MNI = Montreal Neurological Institute coordinates of the peak voxel in the cluster.

brain showing significantly more activity during study of low- than high-imageability pairs. Finally, in line with other studies of word imageability, the imageability effect (greater activity during high- than low-imageability pairs) was not strongly right-lateralized, but if anything, somewhat left-lateralized (Roxbury, McMahon, Coulthard, & Copland, 2016; Pexman, Hargreaves, Edwards, Henry, & Goodyear, 2007; Fiebach & Friederici, 2004). If high-imageability pairs enhance memory via mental imagery, this would suggest that mental imagery in this task is not right-lateralized or bilateral. Alternatively, it may be that imageability does not act via mental imagery, but in some other way, that directly modulates verbal processing and memory, hence the left dominance.

Activity Broken Down by Memory Outcome

Next, we incorporated memory outcome into the analysis, in a 2×2 design: Imageability (high, low) \times Memory (later-recalled, later-forgotten). With our conservative significance criteria ($p < .001$ and $k \geq 19$ voxels; see Methods), this analysis produced numerous significant clusters (Table 2), which we describe next.

Main Effect of Imageability

Only one region in the left superior parietal lobule showed a significant main effect of pair type, with more activity during study of high- than low-imageability pairs (e.g., Wang, Conder, Blitzer, & Shinkareva, 2010; Bedny & Thompson-Schill, 2006; Binder, Westbury, McKiernan, Possing, & Medler, 2005). Because a similar region exhibited an effect of Imageability in the same direction in the previous analysis, this is probably the same effect. The reverse contrast revealed two clusters, in left superior temporal cortex and in right posterior insula. These clusters are close in location to clusters found in the previous analysis, so they may straightforwardly be regions with greater activity during study of low- than high-imageability pairs. The main effects of Imageability thus broadly reinforce the results from the prior section.

Main Effect of Memory

Regions exhibiting greater activity during study of later-remembered than later-forgotten pairs regardless of pair type (main effect of memory without a significant interaction with pair type) included regions that have been found to exhibit subsequent memory effects for verbal materials, including a region similar to Broca's area, in the left ventrolateral pFC, left SMA, linked to covert rehearsal (Koelsch et al., 2009; Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Rushworth, Krams, & Passingham, 2001). Notably, a very large cluster encompassed left putamen, portions of thalamus and brainstem, and a subcluster with a peak in the left hippocampus, converging with the ROI analyses reported above.

The reverse contrast (sometimes referred to as the "subsequent forgetting effect") revealed right TPJ, a major node of the so-called "default-mode network," which has been often reported as being inversely related to memory success (Elman, Rosner, Cohn-Sheehy, Cerreta, & Shimamura, 2013; Turk-Browne, Golomb, & Chun, 2013; Kim, 2011). A region in the right dorsolateral pFC also exhibited more activity during study of subsequently forgotten than recalled pairs. We speculate that this region reflects an effortful, executive function-based strategy that is invoked when a particular pair is difficult to learn but is unsuccessful as a compensatory strategy. Thus, perhaps the right dorsolateral prefrontal activity signals a pair that is challenging to learn but does not provide the participant a way to learn it effectively.

Imageability \times Memory Interactions

Finally, regions outside the hippocampus that show an Imageability \times Memory interaction are candidates for the putative bypassing function. There were two such regions. A cluster in the left caudate nucleus had a larger subsequent memory effect for high- than low-imageability pairs. A right ventrolateral prefrontal region had a significantly greater subsequent memory effect for low- than high-imageability pairs. Because pFC is generally associated with effortful cognitive functions, this may reflect a deliberate strategy, presumably not imagery-dependent, evoked during study of the LL pairs that is effective (in contrast to the right dorsolateral-linked strategy, just mentioned, that showed a subsequent forgetting effect). Thus, the best candidate region for enhancing association-memory in a hippocampal-independent way is the caudate, which seems an unlikely region for this role.

DISCUSSION

Our main findings were that hippocampal activity, primarily left-sided, was greater during study and retrieval of high- than low-imageability pairs, but the subsequent memory effect was not found to differ for high- and low-imageability pairs (Figure 5). The same basic outcome was found for the retrieval success effect (Figure 6). Brain activity outside the hippocampus was, based both on number of significant clusters and voxels, more extensive during study of low- than high-imageability pairs. This supports the Recruitment Hypothesis (Figure 1)—that the imageability effect is due to high-imageability stimuli invoking the hippocampus more than low-imageability stimuli. We found no support for the alternative, Bypassing Hypothesis (Figure 2)—that high imageability improves cued recall by making the task easier by supporting memory in an hippocampal-independent way.

Importantly, this study built on our previous behavioral work on cued recall of word pairs (Madan et al., 2010). That experimental design included all combinations of high- and low-imageability (as well as word frequency) pair

Table 2. Whole-brain Exploratory Analysis of the Study Phase, Incorporating (Subsequent) Memory Outcome, Reporting Clusters Meeting the Criteria $p < .001$ (Single Voxel) and Cluster Size (k) ≥ 19

Cluster No.	MNI (mm)			Peak Voxel Z Score	Volume (No. of Voxels [k])	Region
	x	y	z			
<i>Contrast: High-imageability > Low-imageability</i>						
1	-36	-78	39	3.94	32	L angular gyrus
<i>Contrast: Low-imageability > High-imageability</i>						
2	-48	6	-18	5.13	115	L superior temporal gyrus
3	45	-27	24	4.25	31	R postcentral gyrus
<i>Contrast: Subsequently Remembered > Subsequently Forgotten</i>						
4	-48	15	-12	4.57	226	L superior temporal gyrus/posterior orbital gyrus
5	-6	12	60	4.45	89	L superior frontal gyrus
6	-27	-27	3	4.43	575	(Broken apart with watershed)
6a	-27	-27	3	4.43	173	L putamen
6b	-15	-24	3	4.24	219	L thalamus/brainstem
6c	-33	-15	-9	4.18	142	L hippocampus
6d	-18	3	3	3.83	41	L caudate
7	18	-30	60	4.16	97	R postcentral gyrus
8	-21	-48	60	3.88	24	L posterior parietal cortex
9	-51	-42	27	3.72	27	L supramarginal gyrus
10	-42	-66	12	3.71	27	L intra-occipital sulcus
11	30	-15	6	3.67	132	R putamen
12	-9	57	27	3.67	23	L superior frontal gyrus
<i>Contrast: Subsequently Forgotten > Subsequently Remembered</i>						
13	36	21	39	4.53	71	R dorsolateral prefrontal cortex
14	51	-48	42	4.13	164	R temporoparietal junction
15	-9	-69	33	3.84	34	L precuneus
<i>Contrast, Interaction: [High-remembered - High-forgotten] > [Low-remembered - Low-forgotten]</i>						
16	-33	-6	21	3.68	39	L caudate
<i>Contrast, Interaction: [Low-remembered - Low-forgotten] > [High-remembered - High-forgotten]</i>						
17	48	48	-9	3.58	26	R ventrolateral prefrontal cortex

Clusters 6a–6d are subclusters of Cluster 6 based on the watershed algorithm (see Methods). MNI = Montreal Neurological Institute coordinates of the peak voxel in the cluster.

types, combined with a simple probabilistic model fit to the data to disentangle the effects of stimulus properties on association- versus item-memory. The results pinpointed the effects of imageability to association learning

and could not be explained away by, for example, high-imageability items simply being easier to respond as targets. This laid the groundwork for this study, where we asked whether this boost in association-memory could be

attributed to the hippocampus or to extrahippocampal regions. Moreover, our materials controlled for word frequency, a word property that Madan et al. (2010) found also to increase cued recall but via item-memory and not association-memory effects (see also Criss, Aue, & Smith, 2011). Had we not controlled word frequency, it is reasonable to assume we would have observed a pattern that resembled the Bypassing Hypothesis (Figure 2B, C)—but that would presumably have been because the boost in memory was due to item-memory, already known to be hippocampal independent, rather than association-memory. Second, our stimuli were controlled not only for word length but also for orthographic and phonological neighborhoods, which rules out possible accounts based on competition from other orthographically or phonologically similar stimuli. Furthermore, our high- and low-imageability pairs were matched for semantic similarity. In some sense, a pair of words that are already similar at the start of the experiment implies that such a pair is already known before the experiment. Thus, if high-imageability pairs had been more semantically related than low-imageability pairs, we would have expected that the high-imageability (aka semantically similar) pairs would have been less dependent on hippocampal—or any—activity than the low-imageability (aka semantically less similar) pairs, which would have produced a smaller-sized subsequent memory effect for the high- compared with low-imageability pairs, again misleadingly suggestive of the Bypassing Hypothesis.

Our findings differ in some ways from previous, related studies. Fletcher, Shallice, Frith, Frackowiak, and Dolan (1996) manipulated imageability of word pairs and did not report hippocampal activity (measured with PET) during their vocal cued recall tests; however, they only analyzed brain activity during retrieval blocks and gave low-imageability pairs more study trials in an attempt to cancel out the effect of imageability on accuracy. Fliessbach, Weis, Klaver, Elger, and Weber (2006) reported a between-subject effect; during the retrieval phase of an item recognition task, controlling for word frequency and length, participants with a larger concreteness effect had greater activity in left hippocampus. They did not report any significant within-subject effects in the hippocampus. These findings may have foreshadowed our findings; however, as stated earlier, recognition is typically not hippocampal-dependent. Likewise, Klaver et al. (2005) found an ERP recorded in the hippocampus during continuous recognition to show a main effect of imageability, suggesting that the hippocampus may serve at least to enhance item-memory, but again, recognition memory may not be hippocampal-dependent, and the role of the hippocampus could have been related to noncontrolled properties such as semantic similarity or neighborhood sizes. Murray and Kensinger (2014) asked participants to form interactive imagery (combining two words into a single image) or separation imagery (keeping images of a word pair separate), followed with a surprise associative recognition test. They failed to find hippocampal activity in the univariate analyses of the

task, possibly because of their use of associative recognition or because the inclusion of emotional words may have altered the hippocampal-dependent nature of the task. Duncan, Tompary, and Davachi (2014) did find hippocampal subsequent memory effects in an association-memory task with a variant of associative recognition. Asked to use interactive imagery to learn object pairs, their participants had differential patterns of functional connectivity between hippocampal subregions and other brain areas during encoding versus retrieval. Thus, even associative recognition, with imagery-based strategies, might be hippocampal-dependent. Our findings suggest that, had they included a non-imagery-based strategy as a control condition, they might have found reduced hippocampal activity in that control condition. We also note that there is no indication of a shift from anterior to posterior hippocampus between encoding and retrieval, an idea that has previously been challenged (Schacter & Wagner, 1999).

It is tempting to interpret our effects as relating to participants explicitly applying instructable imagery strategies; however, we suggest our effects differ in important ways from explicitly instructed visual imagery studies. For example, Staresina and Davachi (2010) asked participants to form imagery to remember object-color pairs by imagining the object in the corresponding color. They thought of this as a unitized association, and memory for these unitized associations was less hippocampal-dependent than nonunitized associations. However, even with pairs of high-imageability (concrete) words, participants do not often spontaneously use imagery-based strategies when they are not instructed to do so (Paivio & Yuille, 1969), and our self-reported strategy analyses (not reported here) failed to produce significant effects. Dalton et al. (2016) found right posterior hippocampus to uniquely contribute to nonverbal, visual association-memory (tested with associative recognition); however, our posterior right hippocampal region showed the weakest effects of both imageability and memory in our task; this makes it doubtful that the imageability effect on cued recall derives from hippocampal memory functions specific to visuospatial processes.

The most well-known theory of the superiority of high-imageability words in memory is Paivio's Dual-Coding Theory (Paivio, 1971, 1986). According to Dual-Coding Theory, participants can study high-imageability word pairs with two strategies, forming an image or forming a nonvisual sentence to link the two items; for low-imagery pairs, only the verbal-mediation strategy is available. This implies that the set of regions showing a subsequent memory effect for low-imageability pairs will be a subset of the set of regions showing a subsequent memory effect. Thus, Dual-Coding Theory could be viewed as a specific instantiation of a Bypassing Hypothesis. Instead, we found one region in each interaction contrast. Thus, if an effect related to Dual-Coding Theory was present in our data, it must be quite subtle. Dual-Coding Theory has found some support in neuroimaging data from other tasks, such as lexical decision (Binder et al., 2005) and semantic-similarity

judgments (Sabsevitz, Medler, Seidenberg, & Binder, 2005). The Sabsevitz et al. (2005) study did not report controlling for neighborhood effects or word frequency. However, Binder et al. (2005) did control for these factors. Although not supported for our association-memory task, Dual-Coding Theory may be a good description of other memory and cognitive tasks. On the other hand, our findings converge with other neuroimaging findings inconsistent with Dual-Coding Theory reviewed and reported by Fiebach and Friederici (2004). These authors also cited context availability as a rival theory (Schwanenflugel & Shoben, 1983), but our pools did not differ in contextual diversity (nor semantic similarity or neighborhood effects), making this account unlikely. Dual-Coding Theory also suggests lateralization effects: Whereas low-imageability pairs, amenable only to verbal mediation, should be left-lateralized, high-imageability pairs, amenable to both visual and verbal mediation, should show more bilateral activity. Jessen et al. (2000) found such a pattern during study for an anticipated recognition test. The hippocampus was absent from their findings, which may be in part due to their memory test likely not being hippocampal-dependent (Mayes et al., 2007). However, in an intracranial ERP study, Klaver et al. (2005) found no evidence of lateralization effects in medial-temporal lobe ERPs. It may be that the activity suggestive of Dual-Coding Theory is not essential to memory function and is attenuated when participants are challenged with a difficult task like studying for a subsequent association-memory test.

One behavioral study may have foreshadowed our findings. Peters and Daum (2008) found “remember” responses (intended to measure recollection-based recognition) reduced with age, but more so for concrete (similar to high-imageability) than abstract (similar to low-imageability) words, controlling for word frequency and length. Because recollection is thought to be hippocampal-dependent, they suggested their pattern of findings implied that recollection-based recognition memory is enhanced for concrete words due to hippocampal function; because hippocampal function is reduced with age, it follows that remember responses to concrete words reduce particularly quickly with age. Although indirect, this resonates with our support for the Recruitment Hypothesis, that the imageability advantage in cued recall was associated with greater hippocampal activity in high- than low-imageability pairs.

Finally, in most studies, the more hippocampal-dependent condition is typically, if anything, the more challenging, and thus, the condition with less accuracy; this includes unitization manipulations (e.g., Staresina & Davachi, 2010; Diana et al., 2008; Rhodes & Donaldson, 2008; Quamme et al., 2007; Doeller, Opitz, Krick, Mecklinger, & Reith, 2005), which can be viewed as instantiations of the Bypassing Hypothesis, with few exceptions (Merhav, Karni, & Gilboa, 2015; Bader et al., 2014). This would have led one to predict that low-imageability pairs would have either invoked more hippocampal activity or exhibited a greater depen-

dence (larger magnitude subsequent memory effect) than high-imageability pairs. We found the opposite pattern, which we understand as the imageability of the stimuli recruiting the hippocampus more, ensuring sufficient hippocampal engagement more often during the high- than the low-imageability pairs.

Conclusion

Research is inconclusive with regard to the effect of imageability on association-memory (as well as item-memory, lexical decision, etc.) based on cognitive theories. Our imageability effect cannot be explained away by other stimulus properties that have confounded some other studies. However, our findings suggest why current cognitive accounts of imageability effects on association-memory may not be conclusive: The primary mechanism seems not to be due to additional support or processing by regions outside the hippocampus (Bypassing Hypothesis), but rather, the engagement of the hippocampus (Recruitment Hypothesis). The study of the effects of imageability on association-memory may need to be reformulated: to ask why some materials (such as those high in subjectively rated imageability) engage the hippocampus more than others.

In demonstrating support for the Recruitment Hypothesis, our findings show that the hippocampus can be differentially engaged, depending on stimulus materials. Whereas certain forms of deliberate visual imagery may enable participants to learn and remember associations without relying heavily on the hippocampus, the enhancement of cued recall due to word imageability does the converse, apparently engaging the hippocampus more in a task that is already hippocampal-dependent, and thereby enhancing association-memory. Memory training that is modeled after imageability effects may not increase memory in patient populations with very compromised hippocampal function, as the hippocampal contribution may always be necessary. However, individuals with relatively intact hippocampi, who simply have trouble engaging the hippocampus, may benefit from stimuli with characteristics like those of high-imageability words. Finally, our findings extend boundary conditions of the hippocampal contribution to association-memory to cued recall of word pairs over a range of imageability.

Acknowledgments

We thank Peter Seres, Matthew Brown, and Sanjay Kalra for help designing the fMRI protocol and technical assistance, and Yvonne Chen, Yang Liu, and Isha Ober for assisting with participant testing. We thank Randy McIntosh and Mackenzie Glaholt for assisting with a preliminary experiment that led to this work. Finally, we thank Tobias Sommer, Esther Fujiwara, and Bernhard Staresina for help with interpretation of the results and the manuscript. This study was supported in part by the Natural Sciences and Engineering Research Council of Canada and the Alberta Ingenuity Fund.

Reprint requests should be sent to Jeremy B. Caplan, Department of Psychology, Biological Sciences Building, University of Alberta, Edmonton, Alberta T6G 2E9, Canada, or via e-mail: jcaplan@ualberta.ca.

Notes

1. These results were robust to the selection of participants; when all 31 participants were included, the main effect of Imageability and the Hemisphere \times Imageability interaction were still significant.

2. These results were robust to the specific selection of participants; when all participants with at least one recalled and one forgotten pair of each type (permitting subsequent memory effect and retrieval success effect analyses; $n = 27$) were included, the main effect of Memory was still significant, although the interactions Long Axis \times Memory and Hemisphere \times Long Axis \times Memory were both no longer significant. Imageability still failed to reach significance as a main effect and in all its interactions. Most relevant to the Recruitment Hypothesis, all interactions involving both Memory and Imageability were not significant ($ps > .2$).

REFERENCES

- Bader, R., Opitz, B., Reith, W., & Mecklinger, A. (2014). Is a novel conceptual unit more than the sum of its parts?: fMRI evidence from an associative recognition memory study. *Neuropsychologia*, *61*, 123–134.
- Bedny, M., & Thompson-Schill, S. L. (2006). Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. *Brain and Language*, *98*, 127–139.
- Bell, B. D., & Davies, K. G. (1998). Anterior temporal lobectomy, hippocampal sclerosis, and memory: Recent neuropsychological findings. *Neuropsychology Review*, *8*, 25–41.
- Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, *17*, 905–917.
- Bonner, M. F., Price, A. R., Peelle, J. E., & Grossman, M. (2016). Semantics of the visual environment encoded in parahippocampal cortex. *Journal of Cognitive Neuroscience*, *28*, 361–378.
- Bosch, S. E., Jehee, J. F. M., Fernández, G., & Doeller, C. F. (2014). Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *Journal of Neuroscience*, *34*, 7493–7500.
- Brybaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990.
- Caplan, J. B., Boulton, K. L., & Gagné, C. L. (2014). Associative asymmetry of compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1163–1171.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*, 162–173.
- Criss, A. H., Aue, W. R., & Smith, L. (2011). The effects of word frequency and context variability in cued recall. *Journal of Memory and Language*, *64*, 119–132.
- Dalton, M. A., Hornberger, M., & Piguet, O. (2016). Material specific lateralization of medial temporal lobe function: An fMRI investigation. *Human Brain Mapping*, *37*, 933–941.
- Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 730–740.
- Doeller, C. F., Opitz, B., Krick, C. M., Mecklinger, A., & Reith, W. (2005). Prefrontal-hippocampal dynamics involved in learning regularities across episodes. *Cerebral Cortex*, *15*, 1123–1133.
- Duncan, K., Tomparly, A., & Davachi, L. (2014). Associative encoding and retrieval are predicted by functional connectivity in distinct hippocampal area CA1 pathways. *Journal of Neuroscience*, *34*, 11188–11198.
- Duvernoy, H. M. (1991). *The human brain: Surface, three-dimensional sectional anatomy and MRI*. New York: Springer.
- Elman, J. A., Rosner, Z. A., Cohn-Sheehy, B. I., Cerreta, A. G., & Shimamura, A. P. (2013). Dynamic changes in parietal activation during encoding: Implications for human learning and memory. *Neuroimage*, *82*, 44–52.
- Epp, J., Keith, J. R., Spanswick, S. C., Stone, J. C., Prusky, G. T., & Sutherland, R. J. (2008). Retrograde amnesia for visual memories after hippocampal damage in rats. *Learning & Memory*, *15*, 214–221.
- Fiebach, C. J., & Friederici, A. D. (2004). Processing concrete words: fMRI evidence against a specific right-hemisphere involvement. *Neuropsychologia*, *42*, 62–70.
- Fletcher, P. C., Shallice, T., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1996). Brain activity during memory retrieval: The influence of imagery and semantic cueing. *Brain*, *119*, 1587–1596.
- Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *Neuroimage*, *32*, 1413–1421.
- Geller, A. S., Schleifer, I. K., Sederberg, P. B., Jacobs, J., & Kahana, M. J. (2007). PyEPL: A cross-platform experiment-programming library. *Behavior Research Methods*, *39*, 950–958.
- Gilboa, A., Winocur, G., Grady, C. L., Hevenor, S. J., & Moscovitch, M. (2004). Remembering our past: Functional neuroanatomy of recollection of recent and very remote personal events. *Cerebral Cortex*, *14*, 1214–1225.
- Hammers, A., Allom, R., Koepp, M. J., Free, S. L., Myers, R., Lemieux, L., et al. (2003). Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping*, *19*, 224–247.
- Hannula, D. E., Libby, L. A., Yonelinas, A. P., & Ranganath, C. (2013). Medial temporal lobe contributions to cued retrieval of items and contexts. *Neuropsychologia*, *51*, 2322–2332.
- Henson, R. N. A., Shallice, T., Josephs, O., & Dolan, R. J. (2002). Functional magnetic resonance imaging of proactive interference during spoken cued recall. *Neuroimage*, *17*, 543–558.
- Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*, 673–682.
- Horner, A. J., Bisby, J. A., Bush, D., Lin, W.-J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications*, *6*, 1–11.
- Jackson, O. I., & Schacter, D. L. (2003). Encoding activity in anterior medial temporal lobe supports subsequent associative recognition. *Neuroimage*, *21*, 456–462.
- JASP Team. (2016). JASP (Version 0.7.5.5) [computer software]. Retrieved from jasp-stats.org.
- Jessen, F., Heun, R., Erb, D., Granath, M., Kloese, U., Papassotiropoulos, A., et al. (2000). The concreteness effect: Evidence for dual coding and context availability. *Brain and Language*, *74*, 103–112.

- Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *Neuroimage*, *54*, 2446–2461.
- Klaver, P., Fell, J., Dietl, T., Schür, S., Schaller, C., Elger, C. E., et al. (2005). Word imageability affects the hippocampus in recognition memory. *Hippocampus*, *15*, 704–712.
- Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping*, *30*, 859–873.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.
- Madan, C. R. (2015). Creating 3D visualizations of MRI data: A brief guide. *F1000Research*, *4*, 466.
- Madan, C. R., Glaholt, M. G., & Caplan, J. B. (2010). The influence of item properties on association-memory. *Journal of Memory and Language*, *63*, 46–63.
- Mayer, A., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences*, *11*, 126–135.
- Meltzer, J. A., & Constable, R. T. (2005). Activation of human hippocampal formation reflects success in both encoding and cued recall of paired associates. *Neuroimage*, *24*, 384–397.
- Merhav, M., Karni, A., & Gilboa, A. (2015). Not all declarative memories are created equal: Fast mapping as a direct route to cortical declarative representations. *Neuroimage*, *117*, 80–92.
- Milner, B. (1954). Intellectual function of the temporal lobes. *Psychological Bulletin*, *51*, 42–62.
- Murray, B. D., & Kensinger, E. A. (2013). A review of the neural and behavioral consequences for unitizing emotional and neutral information. *Frontiers in Behavioral Neuroscience*, *7*, 1–12.
- Murray, B. D., & Kensinger, E. A. (2014). The route to an integrative associative memory is influenced by emotion. *PLoS ONE*, *9*, e82372.
- Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, *7*, 217–227.
- Nadel, L., Samsonovich, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. *Hippocampus*, *10*, 352–368.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*, 241–263.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston, Inc.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.
- Paivio, A., & Yuille, J. C. (1969). Changes in associative strategies and paired-associate learning over trials as a function of word imagery and type of learning set. *Journal of Experimental Psychology*, *79*, 458–463.
- Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 words. *Journal of Experimental Psychology*, *76*, 1–25.
- Peters, J., & Daum, I. (2008). Differential effects of normal aging on recollection of concrete and abstract words. *Neuropsychology*, *22*, 255–261.
- Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive Neuroscience*, *19*, 1407–1419.
- Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*, *17*, 230–240.
- Quamme, J. R., Yonelinas, A. P., & Norman, K. A. (2007). Effect of unitization on associative recognition in amnesia. *Hippocampus*, *17*, 192–200.
- Rhodes, S. M., & Donaldson, D. I. (2008). Electrophysiological evidence for the effect of interactive imagery on episodic memory: Encouraging familiarity for non-unitized stimuli during associative recognition. *Neuroimage*, *39*, 873–884.
- Roxbury, T., McMahon, K., Coulthard, A., & Copland, D. A. (2016). An fMRI study of concreteness effects during spoken word recognition in aging. Preservation or attenuation? *Frontiers in Aging Neuroscience*, *7*, 1–12.
- Rushworth, M. F. S., Krams, M., & Passingham, R. E. (2001). The attentional role of the left parietal cortex: The distinct lateralization and localization of motor attention in the human brain. *Journal of Cognitive Neuroscience*, *13*, 698–710.
- Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *Neuroimage*, *27*, 188–200.
- Saling, M. M. (2009). Verbal memory in mesial temporal lobe epilepsy: Beyond material specificity. *Brain*, *132*, 570–582.
- Satterthwaite, T. D., Wolf, D. H., Erus, G., Ruparel, K., Elliott, M. A., Gennatas, E. D., et al. (2013). Functional maturation of the executive system during adolescence. *Journal of Neuroscience*, *33*, 16249–16261.
- Schacter, D. L., & Wagner, A. D. (1999). Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus*, *9*, 7–24.
- Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 82–102.
- Sparks, F. T., Lehmann, H., & Sutherland, R. J. (2011). Between-systems memory interference during retrieval. *European Journal of Neuroscience*, *34*, 780–786.
- Staresina, B. P., & Davachi, L. (2010). Object unitization and associative memory formation are supported by distinct brain regions. *Journal of Neuroscience*, *30*, 9890–9897.
- Stark, C. E. L., & Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 12760–12766.
- Turk-Browne, N. B., Golomb, J. D., & Chun, M. M. (2013). Complementary attentional components of successful memory encoding. *Neuroimage*, *66*, 553–562.
- Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping*, *31*, 1459–1468.
- Wang, W., Yonelinas, A. P., & Ranganath, C. (2013). Dissociable neural correlates of item and context retrieval in the medial temporal lobes. *Behavioural Brain Research*, *254*, 102–107.
- Wolosin, S. M., Zeithamova, D., & Preston, A. R. (2012). Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. *Journal of Cognitive Neuroscience*, *24*, 1532–1547.
- Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *Neuroimage*, *91*, 412–419.