Verbal Paired Associates and the Hippocampus: The Role of Scenes

Ian A. Clark, Misun Kim, and Eleanor A. Maguire

Abstract

■ It is widely agreed that patients with bilateral hippocampal damage are impaired at binding pairs of words together. Consequently, the verbal paired associates (VPA) task has become emblematic of hippocampal function. This VPA deficit is not well understood and is particularly difficult for hippocampal theories with a visuospatial bias to explain (e.g., cognitive map and scene construction theories). Resolving the tension among hippocampal theories concerning the VPA could be important for leveraging a fuller understanding of hippocampal function. Notably, VPA tasks typically use high imagery concrete words and so conflate imagery and binding. To determine why VPA engages the hippocampus, we devised an fMRI encoding task involving closely matched pairs of scene words, pairs of object words, and pairs of very low imagery abstract words. We found

that the anterior hippocampus was engaged during processing of both scene and object word pairs in comparison to abstract word pairs, despite binding occurring in all conditions. This was also the case when just subsequently remembered stimuli were considered. Moreover, for object word pairs, fMRI activity patterns in anterior hippocampus were more similar to those for scene imagery than object imagery. This was especially evident in participants who were high imagery users and not in mid and low imagery users. Overall, our results show that hippocampal engagement during VPA, even when object word pairs are involved, seems to be evoked by scene imagery rather than binding. This may help to resolve the issue that visuospatial hippocampal theories have in accounting for verbal memory. ■

INTRODUCTION

The field of hippocampal neuroscience is characterized by vigorous debates. But one point on which there is general agreement is that people with bilateral hippocampal damage and concomitant amnesia (hippocampal amnesia) are significantly impaired on verbal paired associates (VPA) tasks. The VPA task is a widely used instrument for testing verbal memory and has been a continuous subtest within the Wechsler Memory Scale (WMS) from its initial inception (Wechsler, 1945) to the present day (WMS-IV; Wechsler, 2009). Although the VPA task has been revised many times (e.g., increasing the number of word pairs to be remembered, changing the ratio of difficult to easy word pairs), the basic premise has remained the same. The requirement is to encode pairs of words (e.g., bag–truck), memory for which is then tested. Testing can be conducted in multiple ways, but one primary outcome measure is performance on a delayed cued recall test (i.e., the experimenter asks for the word that goes with bag) 30 min after the completion of the learning trials. Compared with matched healthy control participants, patients with hippocampal amnesia show a consistent and reliable deficit on delayed cued recall tests (Giovanello, Verfaellie, & Keane, 2003; Spiers,

Maguire, & Burgess, 2001; Zola-Morgan, Squire, & Amaral, 1986; Graf & Schacter, 1985), and consequently, the VPA has become emblematic of hippocampal function.

The VPA task is typically regarded as a verbal memory task. However, many theories focus on elucidating the role of the hippocampus in visuospatial rather than verbal processing. This includes accounts that consider spatial navigation (Maguire et al., 2000; O'Keefe & Nadel, 1978), autobiographical memory (Hassabis & Maguire, 2007; Squire, 1992; Scoville & Milner, 1957), scene perception (McCormick, Rosenthal, Miller, & Maguire, 2017; Graham, Barense, & Lee, 2010), the mental construction of visual scene imagery (Zeidman & Maguire, 2016; Maguire & Mullally, 2013), and more specific aspects of visuospatial processing, including perceptual richness, a sense of reliving, and imagery content (St-Laurent, Moscovitch, & McAndrews, 2016; Andrews-Hanna, Reidler, Sepulcre, Poulin, & Buckner, 2010; St. Jacques, Conway, Lowder, & Cabeza, 2010).

The cognitive map theory, for instance, posits that the hippocampus specifically supports flexible, allocentric representations of spatial relationships (O'Keefe & Nadel, 1978). In contrast, the scene construction theory (see also the emergent memory account; Graham et al., 2010) proposes that the anterior hippocampus constructs models of the world in the form of spatially coherent scenes (Dalton & Maguire, 2017; Zeidman & Maguire,

University College London

2016; Maguire & Mullally, 2013; Hassabis & Maguire, 2007). A scene in this context is a specific type of visual image that represents a naturalistic 3-D space typically populated by objects and that is viewed from an egocentric perspective. The construction of scene imagery involves associative processing and binding, but the scene construction theory asserts that the hippocampus is specifically required to perform these functions in the service of creating scene representations (Maguire & Mullally, 2013). The difficulty with theories such as cognitive map and scene construction is that they do not appear to be able to explain why VPA learning is invariably compromised following hippocampal damage.

On the face of it, another hippocampal theory does seem to account for the VPA findings. The relational theory suggests that the hippocampus makes associations between any elements, regardless of whether or not space or scenes are involved (Konkel & Cohen, 2009; Cohen & Eichenbaum, 1993). This generic associative process could account for the creation of an association between two unrelated words in the VPA task, while also explaining the involvement of the hippocampus in visuospatial tasks and the combining of individual elements into a coherent memory or the recombination of different elements from past experiences to simulate the future (Roberts, Schacter, & Addis, 2018; St. Jacques, Carpenter, Szpunar & Schacter, 2018; Thakral, Benoit, & Schacter, 2017; Moscovitch, Cabeza, Winocur, & Nadel, 2016; Schacter et al., 2012). However, a purely associative account of hippocampal function is not completely satisfactory, given that patients with hippocampal damage retain an ability to form associations in some circumstances. For example, intact performance has been reported for Yes/No and forced choice recognition of both intraitem associations and associations between items of the same kind (Mayes et al., 2004), when creating basic associations in probabilistic learning (Kumaran et al., 2007; Knowlton, Mangels, & Squire, 1996), in the rapid learning of arbitrary stimulus–response contingencies (Henson et al., 2017), and in other contexts (see Clark & Maguire, 2016; Mullally & Maguire, 2014).

Resolving the tension among hippocampal theories concerning the VPA could be important for leveraging a fuller understanding of hippocampal function. In taking this issue forward, it is worthwhile first to step back. Examination of the words used in typical VPA tests shows the vast majority are high imagery concrete words. It could be that people use visual imagery when processing the word pairs (Maguire & Mullally, 2013). This speculation has recently received indirect support from the finding that patients with hippocampal amnesia used significantly fewer high imagery words in their narrative descriptions of real and imagined events (Hilverman, Cook, & Duff, 2017), suggesting a potential link between verbal processing and visual imagery.

Currently, therefore, standardized VPA tests may be conflating associative processes and imageability. Patients with hippocampal damage are reportedly unable to imagine fictitious and future scenes in addition to their well-reported memory deficits (Schacter et al., 2012; Race, Keane, & Verfaellie, 2011; Hassabis, Kumaran, Vann, & Maguire, 2007). It would, therefore, follow that their impoverished scene imagery ability may place them at a disadvantage for processing high imagery concrete words. One way to deal with the conflation of visual imagery and binding is to examine very low imagery (abstract) word pairs, which would assess binding outside the realm of imagery. However, abstract word pairs rarely feature in VPA tests used with patients or in neuroimaging experiments.

In addition, different types of high imagery words are not distinguished in VPA tests, with the majority of words representing single objects. However, the scene construction theory links the anterior hippocampus specifically with constructing visual imagery of scenes (Dalton & Maguire, 2017; Zeidman & Maguire, 2016). By contrast, the processing of single objects is usually associated with perirhinal and lateral occipital cortices (Murray, Bussey, & Saksida, 2007; Malach et al., 1995). It could therefore be that a scene word (e.g., forest) in a pair engages the hippocampus (via scene imagery) and not because of binding or visual imagery in general. It has also been suggested that even where each word in a pair denotes an object (e.g., cat–table), this might elicit imagery of both objects together in a scene, and it is the generation of this scene imagery that recruits the hippocampus (Clark & Maguire, 2016; Maguire & Mullally, 2013). Consequently, if visual imagery does play a role in the hippocampal dependence of the VPA task, then it will be important to establish not only whether visual imagery or binding is more relevant but also the type of visual imagery being used.

To determine why VPA engages the hippocampus, we devised an fMRI task with three types of word pairs: where both words in a pair denoted "Scenes," where both words represented single "Objects," and where both words were very low imagery "Abstract" words. This allowed us to separate imageability from binding and to examine different types of imagery. Of particular interest were the Object word pairs because we wanted to ascertain whether they were processed using scene or object imagery. For all word pairs, our main interest was during their initial presentation, when any imagery would likely be evoked.

In addition, we conducted recognition memory tests after scanning to investigate whether the patterns of (hippocampal) activity were affected by whether pairs were successfully encoded or not. Although the VPA memory test used with patients typically involves cued recall, the adaptation of the VPA task for fMRI necessitated the use of recognition memory tests. This is because performing a cued recall test for 135 word pairs that were each seen only once is too difficult even for healthy participants. We note that recognition memory for word pairs is not often tested in patients, and in the few studies where it has been examined, the results are mixed, with some

studies finding a deficit and others a preservation of performance (Mayes et al., 2001; Haist, Shimamura, & Squire, 1992). However, we expected that the use of a recognition memory test instead of cued recall would have limited impact on the patterns of brain activity in this study because we assessed brain activity during the initial presentation of the word pairs and not during memory retrieval. Finally, given that people vary in their use of mental imagery (McAvinue & Robertson, 2007; Kosslyn, Brunn, Cave, & Wallach, 1984; Marks, 1973), we also tested groups of high, mid, and low imagery users to assess whether this influenced hippocampal engagement during VPA encoding.

In line with the scene construction theory, we hypothesized that anterior hippocampal activity would be apparent for Scene words pairs, given the likely evocation of scene imagery. We also predicted that anterior hippocampal activity would be increased for Object word pairs and that this would be best explained by the use of scene imagery. In addition, we expected that the effect of scene imagery use on the hippocampus would be most apparent in high imagery users. By contrast, we predicted that Abstract words pairs would engage areas outside the hippocampus, even when only subsequently remembered pairs were considered.

METHODS

Participants

Forty-five individuals took part in the fMRI study. All were healthy, right-handed, and had normal or corrected-to-normal vision. Given the verbal nature of the task, all participants were highly proficient in English, had English as their first language, and were educated in English throughout their school years. Each participant gave written informed consent. The study was approved by the University College London Research Ethics Committee. Participants were recruited on the basis of their scores on the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). The VVIQ is a widely used self-report questionnaire, which asks participants to bring images to mind and rate them on a 5-point scale as to their vividness (anchored at 1 = *Perfectly clear and as vivid as normal vision* and 5 = *No image at all, you only "know" that you are thinking of the object*). Therefore, a high score on the VVIQ corresponds to low use of visual imagery. The validity of the VVIQ has been demonstrated in numerous ways. For example, experimental studies have found that high visualizers were able to match two pictures more quickly than low visualizers when the first picture had to be retained as a mental image over a 20-sec period (Gur & Hilgard, 1975). In addition, significant correlations between the VVIQ and the Betts' Questionnaire Upon Mental Imagery (another widely used imagery questionnaire; Sheehan, 1967) have also been reported (Campos & Pérez-Fabello, 2005; Burton & Fogarty, 2003).

Our fMRI participants comprised three subgroups ($n = 15$ in each), low imagery users, mid imagery users, and high imagery users. Initially, 184 people completed the VVIQ. Fifteen of the highest and 15 of the lowest scorers made up the low and high imagery groups. A further 15 mid scorers served as the mid imagery group. We acknowledge that these groups are relatively small for an fMRI study, but we were nevertheless interested to see whether any differences would be observed. The groups did not differ significantly on age, gender, years of education, and general intellect. Table 1 provides details of the three groups.

Stimuli

To ensure that any fMRI differences were due to our imagery manipulation and not other word properties, the word conditions were highly matched. Six hundred

Table 1. Characteristics of the Participant Groups

	Imagery Group			p		
	Low	Mid	High	Low vs. Mid	Low vs. High	Mid vs. High
Age, years	23.07 (2.31)	21.87 (2.20)	23.93 (5.26)	.16	.57	.18
No. of male participants	6 (40.0%)	7 (46.67%)	8 (53.33%)	.71	.46	.72
Years of education	16.0 (1.89)	15.8 (1.61)	16.0 (2.33)	.76	1.0	.79
Matrix Reasoning	12.47 (2.26)	11.47 (2.17)	12.07 (3.61)	.23	.72	.59
TOPF	54.93 (5.13)	57.47 (5.49)	53.0 (9.47)	.20	.49	.13
FSIQ	110.13 (5.48)	111.97 (6.13)	110.22 (5.99)	.39	.97	.44
VCI	108.93 (5.32)	110.81 (6.18)	108.75 (6.0)	.38	.93	.36
VVIQ mean score	3.08 (0.45)	2.15 (0.17)	1.51 (0.25)	<.001	<.001	<.001

Means (SDs). Two-tailed p values for t tests (χ^2 test for the number of male participants). General intellect was measured using the Matrix Reasoning subtest (scaled scores) of the Wechsler Adult Intelligence Scale-IV (Wechsler, 2008) and the Test of Premorbid Function (TOPF; Wechsler, 2011), provided an estimate of Full Scale IQ (FSIQ) and a Verbal Comprehension Index (VCI).

fifty-four words were required for the study—218 Scene words, 218 Object words, and 218 Abstract words. Words were initially sourced from databases created by Brysbaert and colleagues, which provided ratings for concreteness, word frequency, age of acquisition, valence, and arousal (Brysbaert, Warriner, & Kuperman, 2014; van Heuven, Mandera, Keuleers, & Brysbaert, 2014; Warriner, Kuperman, & Brysbaert, 2013; Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). It was important to control for valence and arousal given reports of higher emotional ratings for abstract words, which could influence fMRI activity (Vigliocco et al., 2014; Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011). We also used data from the English Lexicon Project (Balota et al., 2007) to provide lexical information about each word—word length, number of phonemes, number of syllables, number of orthographic neighbors, and number of phonological and phonographic neighbors with and without homophones.

To verify that each word induced the expected imagery (i.e., scene imagery, object imagery, or very little/no imagery for the abstract words), we collected two further ratings for each word. First, a rating of imageability to ensure that Scene and Object words were not only concrete but also highly imageable (although concreteness and imageability are often interchanged, and although they are highly related constructs, they are not the same; Paivio, Yuille, & Madigan, 1968) and, additionally, that Abstract words were low on imageability. Second, a decision was elicited about the type of imagery the word brought to mind. This was in response to the following instruction:

> "If you had an image we would like you to classify it as either a 'scene' or an 'object'. A scene is an image in your mind that has a sense of space; that you could step into or operate within. An object on the other hand is more of an isolated image, without additional background imagery. It is also likely that for a number of words you will experience very little or no imagery—please do select this option if this is the case."

These ratings were collected from 119 participants in total using Amazon Mechanical Turk's crowdsourcing Web site, following the procedures used by Brysbaert and colleagues for the databases described above. Words were classified as a Scene or Object word when there was a minimum of 70% agreement on the type of imagery brought to mind, and the mean imageability rating was greater than 3.5 (out of 5). For Abstract words, the mean imageability had to be less than or equal to 2. An overview of the word properties is shown in Table 2. This also includes summary comparison statistics. A list of the words in each category can be found in Appendices A–C and at www.fil.ion.ucl.ac.uk/Maguire/Clark_et_al_2018_Scene_Words.pdf, www.fil.ion.ucl.ac.uk/Maguire/Clark_et_al_2018_Object_Words.pdf, www.fil.ion.ucl.ac.uk/Maguire/Clark_et_al_2018_Abstract_Words.pdf.

Scene, Object, and Abstract words were matched on 13 of the 16 measures. Scene and Object words were matched on all 16 measures, whereas Abstract words, as expected, were less concrete and less imageable than Scene and Object words and had a higher age of acquisition, as is normal for abstract words (Kuperman et al., 2012; Stadthagen-Gonzalez & Davis, 2006). As well as being matched at the overall word type level as shown on Table 2, within each word type, words were assigned to one of four lists (word pairs, single words, catch trials, or postscan memory test lures), and all lists were matched on all measures.

Experimental Design and Task

The fMRI task consisted of two elements, the main task and catch trials. The latter were included to provide an active response element and to encourage concentration during the experiment. To match the WMS-IV Verbal Paired Associate Test (Wechsler, 2009), each stimulus was presented for 4 sec. This was followed by a jittered baseline (a central fixation cross) for between 2 and 5 sec, which aided concentration by reducing the predictability of stimulus presentation (Figure 1D). The scanning session was split into four runs, three runs containing 80 trials lasting 10 min each and a final run of 78 trials lasting 9 min 45 sec. Trials were presented randomly for each participant with no restrictions on what could precede or follow each trial.

Unknown to participants, there were six categories of stimuli—high imagery Scene words, high imagery Object words, and very low imagery Abstract words, shown either in pairs of the same word type (Figure 1A) or as single words (Figure 1B). To equalize visual presentation between the word pairs and the single words, the latter were presented with a random letter string that did not follow the rules of the English language and did not resemble real words (Figure 1B). The average, minimum, and maximum length of the letter strings was matched to the real words. Letter strings could either be presented at the top or the bottom of the screen. There were 45 trials of each condition, with each word shown only once to the participant. Our prime interest was in the word pair conditions and, in particular, the Object word pairs, as these related directly to our research question. The single-word conditions were included for the purposes of specific analyses, which are detailed in the Results section.

Participants were asked to try to commit the real words to memory for later memory tests and were specifically instructed that they would be asked to recall the pairs of real words as pairs. They were explicitly told they would not need to remember the random letter strings. No further instructions about how to memorize the stimuli were given (i.e., we did not tell participants to use any particular strategy).

Participants were told that, occasionally, there would be catch trials where they had to indicate, using a button

Table 2. Properties of Each Word Type

Word Property	Type of Word			p		
	Scene	Object	Abstract	Scene vs. Object	Scene vs. Abstract	Object vs. Abstract
Lexical Criteria						
No. of letters[a]	6.73 (1.99)	6.64 (1.91)	6.75 (1.97)	.64	.90	.55
No. of phonemes[a]	5.61 (1.89)	5.41 (1.71)	5.66 (1.72)	.25	.77	.13
No. of syllables[a]	2.09 (0.85)	2.06 (0.79)	2.16 (0.77)	.68	.38	.18
No. of orthographic neighbors[a]	2.57 (4.97)	3.09 (4.96)	2.35 (4.33)	.28	.62	.10
No. of phonological neighbors[a]	6.28 (11.71)	7.61 (11.80)	6.12 (11.31)	.23	.89	.18
No. of phonological neighbors (including homophones)[a]	6.88 (12.59)	8.13 (12.48)	6.66 (11.86)	.30	.85	.21
No. of phonographic neighbors[a]	1.53 (3.51)	1.78 (3.48)	1.46 (3.21)	.45	.84	.32
No. of phonographic neighbors (including homophones)[a]	1.61 (3.67)	1.96 (3.63)	1.49 (3.36)	.32	.72	.16
Word frequency: Zipf[b]	3.90 (0.71)	3.80 (0.61)	3.88 (0.82)	.12	.77	.27
Age of acquisition[c]	7.69 (2.14)	7.40 (2.12)	9.78 (2.46)	.15	<.001	<.001
Emotional Constructs						
Valence[d]	5.68 (1.08)	5.63 (1.02)	5.58 (1.12)	.63	.34	.61
No. of positive words[d,e]	171 (78.44%)	173 (79.6%)	167 (76.61%)	.81	.65	.49
Hedonic valence[d,f]	1.07 (0.69)	0.98 (0.68)	1.04 (0.70)	.18	.69	.34
Arousal[d]	4.07 (0.96)	3.99 (0.87)	4.04 (0.71)	.34	.73	.46
Imagery						
Concreteness[g]	4.65 (0.22)	4.68 (0.22)	1.83 (0.29)	.11	<.001	<.001
Imageability[h]	4.38 (0.29)	4.41 (0.32)	1.53 (0.20)	.31	<.001	<.001

Means (*SDs*). Two-tailed *p* values for *t* tests (χ^2 test for the number of positive words). Note that each comparison was assessed separately to provide a greater opportunity for any differences between conditions to be identified.

[a]From the English Lexion Project (Balota et al., 2007: exlexicon.wustl.edu).

[b]From van Heuven et al. (2014). The Zipf scale is a standardized measure of word frequency using a logarithmic scale. Values go from 1 (*low-frequency words*) to 6 (*high-frequency words*).

[c]From Kuperman et al. (2012).

[d]From Warriner et al. (2013).

[e]Positive words were those that had a valence score greater than or equal to 5.

[f]Hedonic valence is the distance from neutrality (i.e., from 5), regardless of being positive or negative, as per Vigliocco et al. (2014).

[g]From Brysbaert et al. (2014).

[h]Collected for the current study as detailed in the Methods.

press, if they saw a real word presented with a "pseudo-word" (Figure 1C). The participants were informed that they were not required to remember the real word or the pseudoword presented in these catch trials. A pseudo-word is a combination of letters that resembles a real English word and follows the rules of the English language but is not an actual real word. Pseudowords were gener-ated using the English Lexicon Project (Balota et al., 2007) and were paired with Scene, Object, or Abstract words. They were presented at either the top or the bottom of the screen to ensure that participants attended to both. The number of letters and orthographic neighbors of the pseudowords were matched to all of the real word conditions and across the three pseudoword groups (all
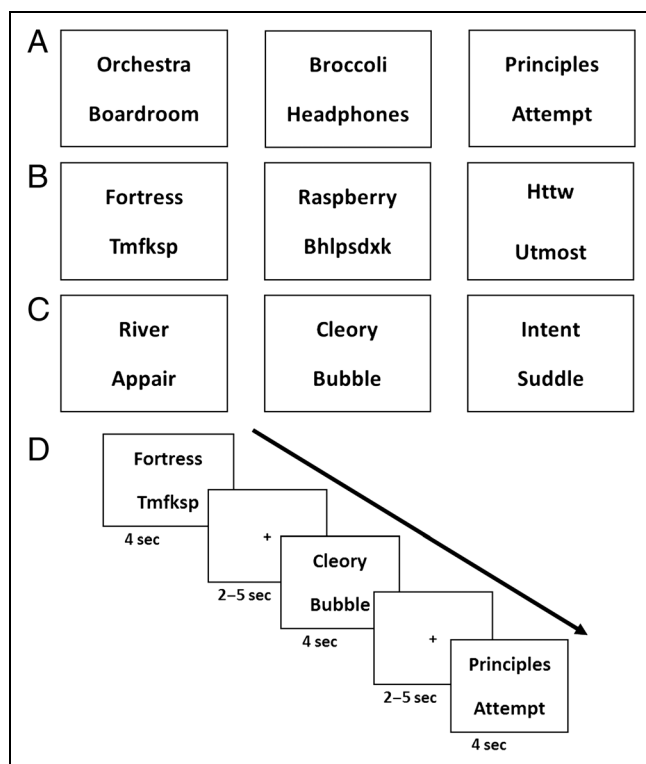
Figure 1. Example stimuli and trial timeline. (A) Examples of stimuli from each of the word types in the order of (from left to right) Scene word pair, Object word pair, and Abstract word pair. (B) Examples of single word trials in the order of (from left to right) Scene single word, Object single word, and Abstract single word. Single words were shown with random letter strings (which could be presented at either the top or the bottom) to be similar to the visual presentation of the word pairs. (C) Examples of catch trials, where a real word was presented with a pseudoword, which could be presented as either the top or bottom word. (D) Example timeline of several trials.

*p*s > .3). In addition, across the pseudoword groups, we matched the accuracy of pseudoword identification (all *p*s > .6) as reported in the English Lexicon Project (Balota et al., 2007). Forty-eight catch trials were presented over the course of the experiment, 16 trials with each of the word types, ranging between 10 and 15 in each of the four runs. Catch trials were pseudorandomly presented to ensure regular presentation, but not in a predictable manner. Feedback was provided at the end of each scanning run as to the number of correctly identified pseudowords and incorrectly identified real words.

Postscan Recognition Memory Tests

Following scanning, participants had two recognition memory tests. The first was an item recognition memory test for all 405 words presented during scanning (45 words for each of three single word types and 90 words for each of three paired word types) and a further 201 foils (67 of each word type). Each word was presented on its own in the center of the screen for up to 5 sec. Words were presented randomly in a different order for each participant. Participants had to indicate for each word whether they had seen it in the scanner (old) or not (new). Following this, they rated their confidence in their answer on a 3-point scale—high confidence, low confidence, or guessing. Any trials where a participant correctly responded "old" and then indicated they were guessing were excluded from subsequent analyses.

After the item memory test, memory for the pairs of words was examined. This associative memory test presented all of the 135 word pairs shown to participants in the scanner and an additional 66 lure pairs (22 of each type), one pair at a time, for up to 5 sec. The word pairs were presented in a different random order for each participant. The lure pairs were constructed from the single words that were presented to the participants in the scanner. Therefore, the participants had seen all of the words presented to them in the associative recognition memory test, but not all were previously in pairs, specifically testing whether the participants could remember the correct associations. Participants were asked to indicate whether they saw that exact word pair presented to them in the scanner (old) or not (new). They were explicitly told that some pairs would be constructed from the single words they had seen during scanning and not to make judgments solely on individual words, but to consider the pair itself. Confidence ratings were obtained in the same way as for the item memory test, and trials where a participant correctly responded "old" and then indicated they were guessing were excluded from subsequent analyses.

Debriefing

On completion of the memory tests, participants were asked about their strategies for processing the words while they were in the scanner. At this point, the participants were told about the three different types of words presented to them—Scenes, Objects, and Abstract. For each word type and separately for single words and word pairs, participants were presented with reminders of the words. They were then asked to choose from a list of options as to which strategy best reflected how they processed that word type. Options included the following: "I had a visual image of a scene related to this type of single word" (scene imagery for single words), "I had a visual image of the two entities that the words represented within a single visual scene" (scene imagery for word pairs), "I had a visual image of a single entity (e.g., one specific object) for a word with no other background imagery" (object imagery), "I read each word without forming any visual imagery at all" (no imagery).

Statistical Analyses of the Behavioral Data

Stimuli Creation and Participant Group Comparisons

Comparisons between word conditions and between the participant groups were performed using independent samples *t* tests for continuous variables and χ^2 tests

for categorical variables. An alpha level of $p > .05$ was used to determine that the stimuli/groups were matched. Note that each comparison was assessed separately (using t tests or χ^2 tests) to provide a greater opportunity for any differences between conditions to be identified.

Main Study

Both within- and between-participant designs were used. The majority of analyses followed a within-participant design, with all participants seeing all word conditions. In addition, participants were split into three groups dependent on their VVIQ score allowing for between-participant analyses to be performed.

All data were assessed for outliers, defined as values that were at least 2 standard deviations away from the mean. If an outlier was identified, then the participant was removed from the analysis in question (and this is explicitly noted in the Results section). Memory performance for each word condition was compared with chance level (50%) using one-sample t tests. For all within-participant analyses, when comparing across three conditions, repeated-measures ANOVAs with follow-up paired t tests were used, and for comparison across two conditions, paired t tests were utilized. For between-participant analyses, a one-way ANOVA was performed with follow-up independent samples t tests.

All ANOVAs were subjected to Greenhouse–Geisser adjustment to the degrees of freedom if Mauchly's sphericity test identified that sphericity had been violated. For all statistical tests, alpha was set at .05. Effect sizes are reported following significant results as Cohen's d for one sample and independent sample t tests, eta-squared for repeated-measures ANOVA, and Cohen's d for repeated measures (d_{rm}) for paired samples t tests (Lakens, 2013). All analyses were performed in IBM SPSS statistics v22.

Scanning Parameters and Data Preprocessing

T2*-weighted echo-planar images (EPI) were acquired using a 3T Siemens Trio scanner (Siemens Healthcare) with a 32-channel head coil. fMRI data were acquired over four scanning runs using scanning parameters optimized for reducing susceptibility-induced signal loss in the medial-temporal lobe: 48 transverse slices angled at $-30°$, repetition time = 3.36 sec, echo time (TE) = 30 msec, resolution = $3 \times 3 \times 3$ mm, matrix size = 64×74, z-shim gradient moment of -0.4 mT/m msec (Weiskopf, Hutton, Josephs, & Deichmann, 2006). Fieldmaps were acquired with a standard manufacturer's double-echo gradient-echo field map sequence (short TE = 10 msec, long TE = 12.46 msec, 64 axial slices with 2-mm thickness and 1-mm gap yielding whole-brain coverage; in-plane resolution 3×3 mm). After the functional scans, a 3D MDEFT structural scan was obtained with 1-mm isotropic resolution (Deichmann, Schwarzbauer, & Turner, 2004).

Preprocessing of data was performed using SPM12 (www.fil.ion.ucl.ac.uk/spm). The output of the SPM image realignment protocol showed that head motion was low (mean [SD] in millimeters: $x = 0.51$ [0.32], $y = 1.29$ [0.33], $z = 1.64$ [0.80]; mean [SD] in degrees: pitch = 0.03 [0.03], roll = 0.01 [0.01], yaw = 0.01 [0.01]) and was smaller than the voxel size. Functional images were co-registered to the structural image and then realigned and unwarped using field maps. The participant's structural image was segmented and spatially normalized to a standard EPI template in MNI space with a voxel size of $2 \times 2 \times 2$ mm and the normalization parameters were then applied to the functional data. For the univariate analyses, the functional data were smoothed using an 8-mm FWHM Gaussian kernel. In line with published representational similarity analysis (RSA) literature (e.g., Chadwick, Jolly, Amos, Hassabis, & Spiers, 2015; Marchette, Vass, Ryan, & Epstein, 2014; Kriegeskorte, Mur, Ruff, et al., 2008), the multivariate analyses used unsmoothed data. We used unsmoothed data to capture neural information in the form of spatially distributed activity across multiple voxels. Smoothing potentially washes out the fine activity differences between voxels.

Where bilateral ROI analyses were performed, the hippocampal ROIs were manually delineated on a previously collected ($n = 36$) group-averaged structural MRI scan ($1 \times 1 \times 1$ mm) using ITK-SNAP (www.itksnap.org) and then resampled to our functional scans ($2 \times 2 \times 2$ mm). The anterior hippocampus was delineated using an anatomical mask that was defined in the coronal plane and went from the first slice where the hippocampus can be observed in its most anterior extent until the final slice of the uncus. In terms of structural space, this amounted to 3616 voxels and in functional space to 481 voxels. The posterior hippocampus was defined as proceeding from the first slice following the uncus until the final slice of observation in its most posterior extent (see Dalton, Zeidman, Barry, Williams, & Maguire, 2017, for more details). In terms of structural space, this amounted to 4779 voxels and in functional space to 575. The whole hippocampus mask combined the anterior and posterior masks and therefore contained 8395 voxels in structural space and 1056 voxels in functional space.

fMRI Analysis: Univariate

The six experimental word conditions were Scene, Object, and Abstract words, presented as either word pairs or single words. As noted above, our prime interest was in the word pair conditions and, in particular, the Object word pairs, as these related directly to our research question. We therefore directly contrasted fMRI BOLD responses between the word pair conditions. The single-word conditions were included for the purposes of specific analyses, which are detailed in the Results section. We performed two types of whole-brain analysis, one using all of the trials (45 per condition) and the other using only trials where the

items were subsequently remembered, not including trials where the participant indicated they were guessing. The average number of trials per condition were as follows: Scene word pairs, 31.49 ($SD = 6.25$); Object word pairs, 34.53 ($SD = 5.84$); Abstract word pairs, 27.42 ($SD = 8.67$). See Table 4 for comparisons of the number of correct trials, not including guessing, across the conditions.

For both analyses, the general linear model consisted of the word condition regressors convolved with the hemodynamic response function, in addition to participant-specific movement regressors and physiological noise regressors. The Artifact Detection Toolbox (www.nitrc.org/projects/artifact_detect/) was used to identify spikes in global brain activation, and these were entered as a separate regressor. Participant-specific parameter estimates for each regressor of interest were calculated for each voxel. Second-level random effects analyses were then performed using one-sample t tests on the parameter estimates. For comparison across VVIQ imagery groups, we performed an ANOVA with follow-up independent sample t tests. We report results at a peak-level threshold of p less than .001 whole-brain uncorrected for our a priori ROI—the hippocampus—and p less than .05 family-wise error (FWE)-corrected at the voxel level elsewhere.

In addition, several ROI analyses were performed on a subset of the univariate analyses. Three ROIs were considered—the whole hippocampus, the anterior hippocampus, and the posterior hippocampus (all bilateral). We used a peak-level threshold of p less than .05 FWE-corrected at the voxel level for each mask and, where indicated in the Results section, also a more lenient threshold of p less than .001 uncorrected for each mask.

fMRI Analysis: Multivariate

Multivoxel pattern analysis was used to test whether the neural representations of the Object word pairs were more similar to the Scene single words than the Object single words when separately examining bilateral anterior and posterior hippocampal ROIs. For each participant, t statistics for each voxel in the ROI were computed for each condition (Object word pair, Object single word, Scene single word) and in each scanning run. The Pearson correlation between each condition was then calculated as a similarity measure (Object word pair/Object word pair, Object word pair/Scene single word, Object word pair/Object single word). The similarity measure was cross-validated across different scanning runs to guarantee the independence of each data set. Repeated-measures ANOVA and paired t tests were used to compare the similarity between conditions at the group level. This multivariate analysis was first applied to the data from all participants and then to the three subsets of participants (low, mid, and high imagery users). All data were assessed for outliers, defined as values that were at least 2 standard deviations away from the group mean. If an outlier was identified, then the participant was removed from the analysis in question (and this is explicitly noted in the Results section).

Note that the absolute correlation of the similarity value is expected to be low because of inherent neural variability and the fact that a unique set of words was presented for each scanning run. As such, the important measure is the comparison of the similarity value between the conditions, not the absolute similarity value of a single condition. The range of similarity values that we found was entirely consistent with those reported in other studies utilizing a similar representational similarity approach in a variety of learning, memory, and navigation tasks in a wide range of brain regions (Kim, Jeffery, & Maguire, 2017; Bellmund, Deuker, Navarro Schröder, & Doeller, 2016; Deuker, Bellmund, Navarro Schröder, & Doeller, 2016; Schapiro, Turk-Browne, Norman, & Botvinick, 2016; Schuck, Cai, Wilson, & Niv, 2016; Chadwick et al., 2015; Hsieh & Ranganath, 2015; Milivojevic, Vicente-Grabovetsky, & Doeller, 2015; Hsieh, Gruber, Jenkins, & Ranganath, 2014; Staresina, Henson, Kriegeskorte, & Alink, 2012).

RESULTS

Behavioral

On average, participants identified 85.56% ($SD = 11.52$) of the pseudowords during catch trials, showing that they maintained concentration during the fMRI experiment. On the postscan item memory test, Scene, Object, and Abstract words were remembered above chance, and there were no differences between the conditions (Table 3, which includes the statistics). Performance on the associative memory test also showed that Scene, Object, and Abstract word pairs were remembered above chance (Table 4, which includes the statistics). Considering the average performance across the four word conditions used in the main univariate analyses (i.e., Scene word pairs, Object word pairs, Abstract word pairs, Abstract single words) then one participant performed below chance. The fMRI analyses do not change whether this participant is included or not. Comparison of memory performance across the word types found differences in performance in line with the literature (Paivio, 1969). Both types of high imagery word pairs (Scene and Object) were remembered better than Abstract word pairs (Figure 2; Table 4), whereas Object word pairs were remembered better than Scene word pairs. Given that the word pair memory lures were highly confusable with the actual word pairs (because the lure pairs were made up of the studied single words), d' values were also calculated for the word pairs. Scene, Object, and Abstract word pairs all showed d' values greater than 0, representing the ability to discriminate between old and new pairs (Table 4). Both Scene and Object word pairs had greater d' values than Abstract word pairs, and Object word pairs d' values were greater

Table 3. Performance (% Correct) on the Postscan Item Memory Test (Nonguessing Trials)

	Scene Single Words	Object Single Words	Abstract Single Words
Mean	67.41	66.37	67.61
SD	14.93	17.71	16.06

	Comparison to Chance (50%)		
	t(44)	p	d
Scene single words	7.82	<.001	2.36
Object single words	6.20	<.001	1.87
Abstract single words	7.36	<.001	2.22

	Comparison across the Word Types	
	F(1.76, 77.51)	p
Main effect	0.28	.73

than those for Scene word pairs (Table 4), showing the same pattern as that calculated using the percentage correct. Overall, these behavioral findings show that, despite the challenging nature of the experiment with so many stimuli, participants engaged with the task committed a good deal of information to memory and could successfully distinguish between previously presented word pairs and highly confusable lures.

fMRI

Univariate Analyses

We performed two whole-brain analyses, one using all of the trials and another using only trials where the items were subsequently remembered in the postscan memory tests (the item memory test for the single word trials, the associative memory test for the word pairs, excluding trials where participants correctly responded "old" and then indicated that they were guessing). The two analyses yielded very similar results across the whole brain, even though the analysis using only subsequently remembered stimuli was less well powered because of the reduced number of stimuli. Given that our interest was in the point at which participants were initially processing the word pairs and potentially using mental imagery to do so, we focus on the results of the analysis using all of

Table 4. Performance (% Correct and d') on the Postscan Associative Memory Test (Nonguessing Trials)

	Percent Correct			d'		
	Scene Word Pairs	Object Word Pairs	Abstract Word Pairs	Scene Word Pairs	Object Word Pairs	Abstract Word Pairs
Mean	69.98	76.74	60.94	1.07	1.33	0.74
SD	13.88	12.97	19.27	0.81	0.83	0.56

	Comparison to Chance (50%)			Comparison to Chance (0)		
	t(44)	p	d	t(44)	p	d
Scene word pairs	9.65	<.001	2.91	8.83	<.001	2.66
Object word pairs	13.83	<.001	4.17	10.84	<.001	3.27
Abstract word pairs	3.81	<.001	1.15	8.94	<.001	2.70

	Comparison across the Word Types			Comparison across the Word Types		
	F(1.35, 59.48)	p	η^2	F(2, 88)	p	η^2
Main effect	24.21	<.001	.36	23.75	<.001	.35

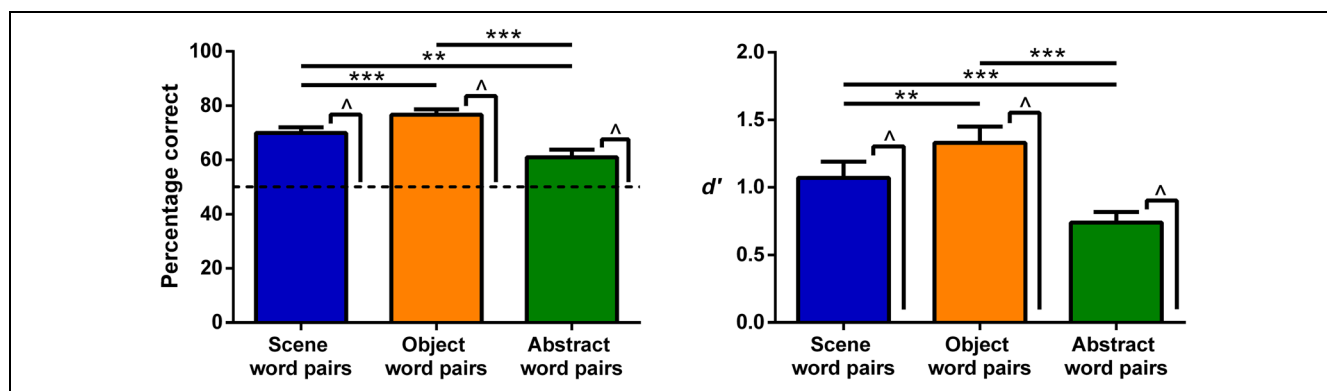|	t(44)	p	d_{rm}	t(44)	p	d_{rm}
Scene vs. Object	5.25	<.001	0.50	3.06	.004	0.32
Scene vs. Abstract	3.58	.001	0.52	4.35	<.001	0.42
Object vs. Abstract	5.75	<.001	0.94	6.22	<.001	0.78

Figure 2. Memory performance on the associative memory test shown by percentage correct (left) and d' (right). Error bars are 1 *SEM*. ^ indicates a significant difference from chance (for percentage correct the dashed line indicates chance at 50%, for d' it is 0) at $p < .001$. Stars show the significant differences across the word pair types: $**p < .01$, $***p < .001$.

the trials. Results are also reported for the analyses using only the remembered stimuli, which allowed us to control for any memory-related effects.

We first compared the high imagery (Scene, Object) and very low imagery (Abstract) word pairs. All of the conditions involved associative processing, and so we reasoned that any differences we observed, particularly in hippocampal engagement, would be due to the imageability of the Scene and Object word pairs. As predicted, Scene word

pairs elicited greater bilateral anterior (but not posterior) hippocampal activity compared with Abstract word pairs (Figure 3A; see full details in Table 5A). Of note, increased activity was also observed in bilateral parahippocampal, fusiform, retrosplenial, and left ventromedial prefrontal cortices. The analysis using only the remembered stimuli showed very similar results, including for the anterior hippocampus (Table 6A). The reverse contrast identified no hippocampal engagement, but rather greater activity for
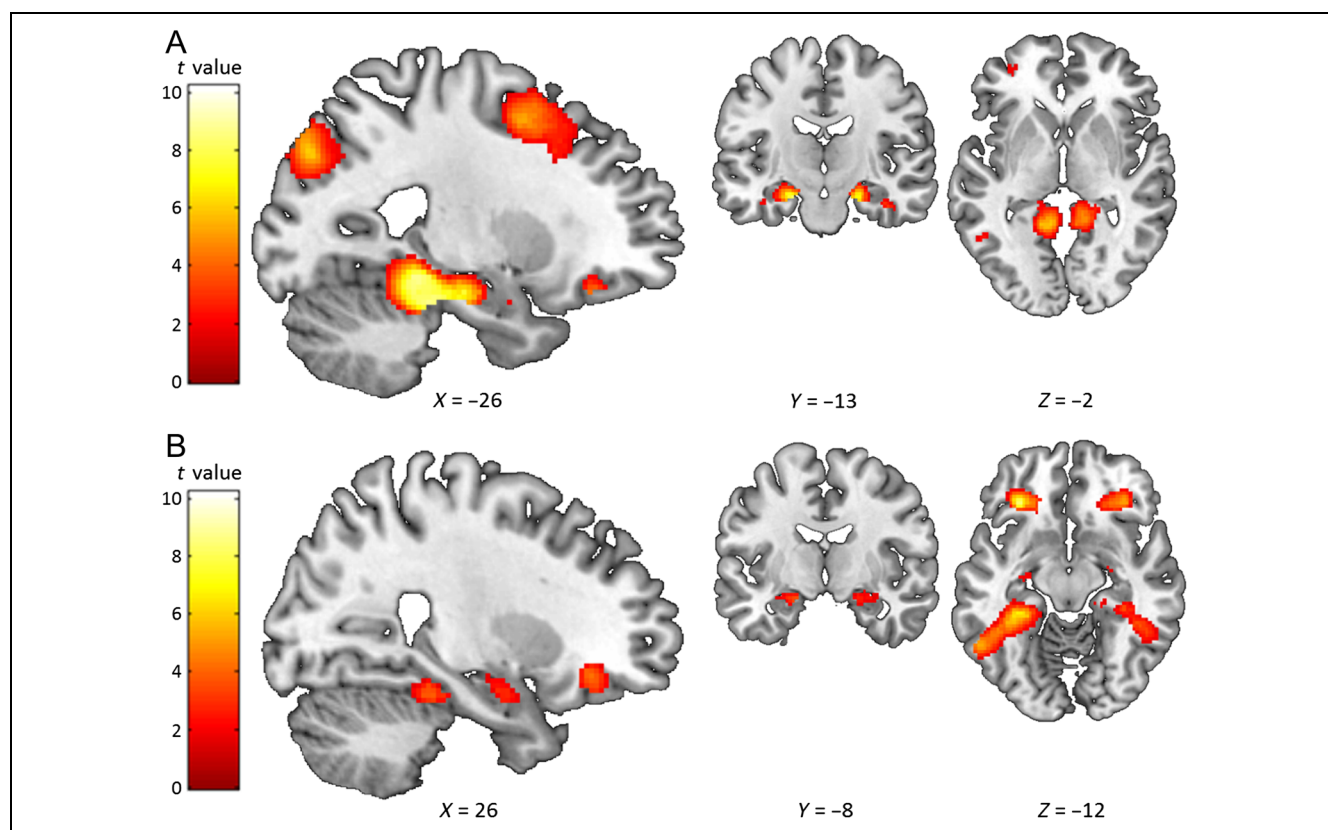

Figure 3. Comparison of high imagery Scene or Object word pairs with very low imagery Abstract word pairs. The sagittal slice is of the left hemisphere, which is from the ch2better template brain in MRicron (Rorden & Brett, 2000; Holmes et al., 1998). The left of the image is the left side of the brain. The colored bar indicates the t value associated with each voxel. (A) Scene word pairs > Abstract word pairs. (B) Object word pairs > Abstract word pairs. Images are thresholded at $p < .001$ uncorrected for display purposes.

Table 5. High Imagery Word Pairs Compared with Abstract Word Pairs

Region	Peak Voxel Coordinates	t
A. Scene Word Pairs > Abstract Word Pairs		
Left anterior hippocampus	−20, −16, −20	8.79
Right anterior hippocampus	20, −10, −20	7.58
Left retrosplenial cortex	−10, −52, 4	9.15
Left fusiform cortex	−22, −34, −20	9.03
Right retrosplenial cortex	10, −48, 6	8.73
Left middle occipital cortex	−30, −74, 34	8.49
Right parahippocampal cortex	24, −34, −20	8.40
Left inferior temporal cortex	−56, −54, −10	8.03
Right fusiform cortex	32, −32, −14	7.72
Left ventrolateral prefrontal cortex	−30, 32, −16	6.51
Right middle occipital cortex	44, −70, 26	6.48
Left middle frontal cortex	−26, 6, 50	6.19
Left inferior frontal cortex	−42, 32, 12	5.74
B. Object Word Pairs > Abstract Word Pairs		
Left anterior hippocampus	−20, −10, −18	4.45
Right anterior hippocampus	20, −10, −18	3.98
Left ventral medial prefrontal cortex	−32, 32, −14	9.45
Left fusiform cortex (extending to parahippocampal cortex)	−32, −34, −20	8.88
Left middle occipital cortex	−34, −80, 28	6.17
Right ventrolateral prefrontal cortex	34, 32, −12	6.05
Left inferior frontal cortex	−40, 28, 14	6.05
Right fusiform gyrus (extending to parahippocampal cortex)	34, −32, −18	5.72

$p < .001$ uncorrected for the hippocampus and $p < .05$ FWE-corrected for the rest of the brain. Brain regions within the medial temporal lobe were identified via visual inspection. For regions outside the medial temporal lobe, the AAL atlas was used (Tzourio-Mazoyer et al., 2002).

Abstract word pairs in middle temporal cortex (−58, −36, −2, $T = 6.58$) and temporal pole (−52, 10, −22, $T = 6.16$).

Object word pairs also showed greater bilateral anterior (but not posterior) hippocampal activity compared with the Abstract word pairs, along with engagement of bilateral parahippocampal cortex, fusiform cortex, and ventromedial prefrontal cortex (Figure 3B; Table 5B), with increased anterior hippocampal activity also apparent when just the subsequently-remembered stimuli were considered (Table 6B). The reverse contrast identified no

hippocampal engagement, but rather greater activity for Abstract word pairs in middle temporal cortex (−62, −32, −2, $T = 8$) and temporal pole (−54, 10, −18, $T = 7.12$).

Increased anterior hippocampal activity was therefore observed for both Scene and Object word pairs compared with the very low imagery Abstract word pairs. As greater anterior hippocampal engagement was apparent even when using just the remembered stimuli, it is unlikely that this result can be explained by better associative memory or successful encoding for the high imagery word pairs. Rather the results suggest that the anterior hippocampal activity for word pair processing may be related to the use of visual imagery.

All of the above contrasts involved word pairs, suggesting that associative binding per se cannot explain the

Table 6. Remembered High Imagery Word Pairs Compared with Remembered Abstract Word Pairs

Region	Peak Voxel Coordinates	t
A. Scene Word Pairs Remembered > Abstract Word Pairs Remembered		
Left hippocampus	−28, −22, −18	7.53
Right hippocampus	24, −20, −18	5.09
Left retrosplenial cortex	−10, −50, 2	10.20
Left fusiform cortex (extending to parahippocampal cortex)	−30, −34, −14	8.21
Right retrosplenial cortex	10, −48, 4	7.47
Left middle occipital lobe	−30, −80, 40	7.23
Right fusiform cortex (extending to parahippocampal cortex)	26, −28, −20	7.00
Left ventral medial prefrontal cortex	−30, 34, −12	6.51
Right middle occipital lobe	44, −70, 28	6.29
Left inferior temporal cortex	−56, −54, −10	5.64
B. Object Word Pairs Remembered > Abstract Word Pairs Remembered		
Left hippocampus	−32, −22, −12	5.05
Left ventral medial prefrontal cortex	−30, 34, −12	9.26
Left fusiform cortex (extending to parahippocampal cortex)	−30, −32, −18	7.94
Left middle occipital lobe	−34, −82, 30	6.43
Left inferior temporal cortex	−54, −58, −6	6.13

$p < .001$ uncorrected for the hippocampus and $p < .05$ FWE-corrected for the rest of the brain. Brain regions within the medial temporal lobe were identified via visual inspection. For regions outside the medial temporal lobe, the AAL atlas was used (Tzourio-Mazoyer et al., 2002).

results. However, it could still be the case that binding Abstract word pairs does elicit increased hippocampal activity but at a lower level than Scene and Object word pairs. To address this point, we compared the Abstract word pairs with the Abstract single words, as this should reveal any hippocampal activity related to associative processing of the pairs. No hippocampal engagement was evident for the Abstract word pairs in comparison to the Abstract single words (Table 7). This was also the case when just the remembered stimuli were considered (Table 8), albeit with slightly lower power than the previous contrasts (number of trials for the Abstract word pairs = 27.42 [SD = 8.67]; for the Abstract single words: 30.42 [SD = 7.23]).

Given the difficulty of interpreting null results, in particular when using whole-brain standard contrasts, we performed additional ROI analyses to further test whether any subthreshold hippocampal activity was evident for the Abstract word pairs compared with the Abstract single words. Using an anatomically defined bilateral whole hippocampal mask, no differences in hippocampal activity were apparent at a $p < .05$ FWE-corrected threshold at the voxel level for the mask or when a more lenient $p < .001$ uncorrected threshold was used. We then extracted average beta values from across the whole hippocampus bilateral ROI and two additional smaller ROIs—anterior and posterior hippocampus—for the Abstract word pairs and Abstract single words. t Tests showed that there were no differences between conditions (whole hippocampus: $t(44) = 0.16, p = .88$; anterior hippocampus only: $t(44) = 0.13, p = .89$; posterior hippocampus only: $t(44) = 0.18, p = .86$). Similar results were also observed when using just the remembered stimuli (whole hippocampus: $t(44) = 1.16, p = .25$; anterior hippocampus only: $t(44) = 1.36, p = .18$; posterior hippocampus

Table 7. Abstract Word Pairs Compared with Abstract Single Words

Region	Peak Voxel Coordinates	t
Abstract Word Pairs > Abstract Single Words		
Left middle temporal cortex	−64, −36, 2	8.39
Left temporal pole	−52, 12, −16	6.72
Left fusiform cortex	−38, −46, −20	6.64
Left inferior frontal cortex	−54, 24, 12	6.54
Left inferior occipital cortex	−42, −68, −12	6.52
Right inferior occipital cortex	36, −74, −12	6.11
Right lingual cortex	20, −82, −10	5.87
Left precentral gyrus	−50, 0, 48	5.84

$p < .001$ uncorrected for the hippocampus (no activations found) and $p < .05$ FWE-corrected for the rest of the brain. Brain regions within the medial temporal lobe were identified via visual inspection. For regions outside the medial temporal lobe, the AAL atlas was used (Tzourio-Mazoyer et al., 2002).

Table 8. Remembered Abstract Word Pairs Compared with Remembered Abstract Single Words

Region	Peak Voxel Coordinates	t
Abstract Word Pairs Remembered > Abstract Single Words Remembered		
Left inferior frontal gyrus	−54, 14, 12	9.50
Left precentral gyrus	−48, −2, 48	8.02
Left middle temporal gyrus	−52, −46, 4	8.21
Left inferior occipital lobe	−38, −78, −8	7.23
Right inferior occipital lobe	34, −80, −6	7.11
Left supplementary motor area	−2, 4, 56	6.72
Right inferior frontal gyrus	50, 10, 28	6.44
Right superior temporal pole	46, −30, 4	6.11
Right caudate nucleus	12, 10, 6	6.07
Left pallidum	−18, 6, 0	6.04

$p < .001$ uncorrected for the hippocampus (no activations found) and $p < .05$ FWE-corrected for the rest of the brain. Brain regions within the medial temporal lobe were identified via visual inspection. For regions outside the medial temporal lobe, the AAL atlas was used (Tzourio-Mazoyer et al., 2002).

only: $t(44) = 0.63, p = .53$). Overall, therefore, even at lenient thresholds and using an ROI approach, no hippocampal engagement was identified for Abstract word pairs compared with the Abstract single words. Although the absence of evidence is not evidence of absence, this is in direct contrast to our findings of increased hippocampal activity for the high imagery word pairs compared with the very low imagery Abstract word pairs. This, therefore, lends support to the idea that the use of visual imagery might be important for inducing hippocampal responses to word pairs.

We also predicted that anterior hippocampal activity would be specifically influenced by the use of scene imagery, as opposed to visual imagery per se. The inclusion of both Scene and Object word pairs offered the opportunity to test this. Scene word pairs would be expected to consistently evoke scene imagery (as both words in a pair represented scenes), whereas Object word pairs could evoke both or either object and scene imagery (e.g., object imagery by imagining the two objects without a background context or scene imagery by creating a scene and placing the two objects into it), thus potentially diluting the hippocampal scene effect. Scene word pairs might therefore activate the anterior hippocampus to a greater extent than Object word pairs. This comparison also provided an additional opportunity to contrast the effects of scene imagery and memory performance on hippocampal activity, because Object word pairs were better remembered than the Scene word pairs. As such, if hippocampal activity could be better explained by word pair memory performance
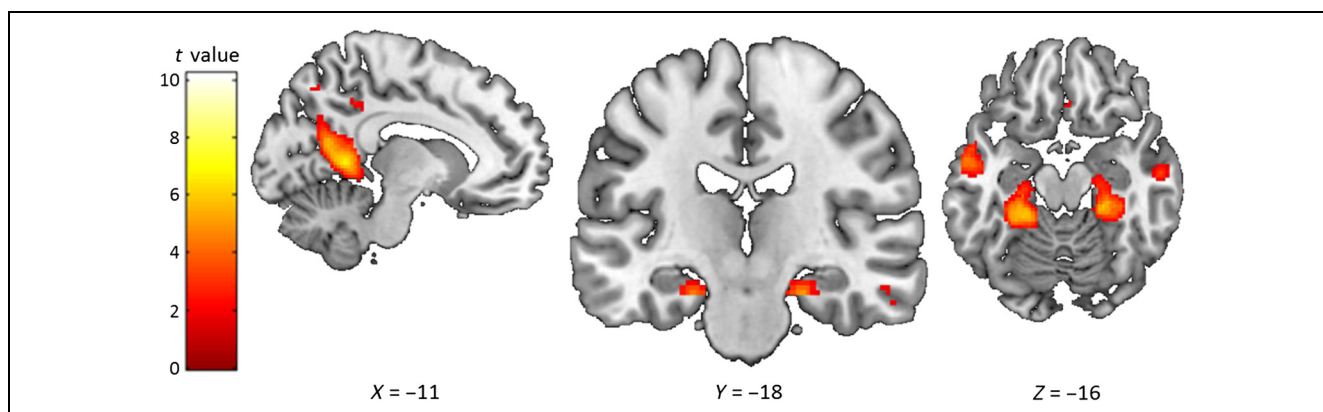
Figure 4. Brain areas more activated by Scene word pairs than Object word pairs. The sagittal slice is of the left hemisphere, which is from the ch2better template brain in MRicron (Rorden & Brett, 2000; Holmes et al., 1998). The left of the image is the left side of the brain. The colored bar indicates the *t* value associated with each voxel. Images are thresholded at $p < .001$ uncorrected for display purposes.

rather than scene imagery, we would expect that Object word pairs would show greater hippocampal activity than Scene word pairs.

Contrasting Scene and Object word pairs revealed that, in line with our prediction, Scene word pairs evoked greater bilateral anterior (but not posterior) hippocampal activity than the Object word pairs (Figure 4; Table 9A). Analysis using just the remembered stimuli gave similar results (Table 10A). Other areas that showed increased activity for the Scene word pairs included the retrosplenial and parahippocampal cortices. The reverse contrast examining what was more activated for Object word pairs compared with Scene word pairs found no evidence of hippocampal activity despite better subsequent memory performance for the Object word pairs (Table 9B), even when just the remembered stimuli were examined (Table 10B). It seems, therefore, that the anterior hippocampus may be particularly responsive to scene imagery and that increases in hippocampal activity in this task were not driven by greater memory performance.

To summarize, our univariate analyses found that Scene word pairs engaged the anterior hippocampus the most, followed by the Object word pairs, with the Abstract word pairs not eliciting any significant increase in activation (Figure 5). This is what we predicted and may be suggestive of particular responsivity of the anterior hippocampus to scenes.

Multivariate Analyses

We next sought further, more direct evidence that our main condition of interest, Object word pairs, elicited hippocampal activity via scene imagery. Given our univariate findings of increased anterior hippocampal activity for Scene word pairs and Object word pairs compared with Abstract word pairs and the extant literature showing the importance of the anterior hippocampus for processing scenes (e.g., Zeidman & Maguire, 2016, but see also Sheldon & Levine, 2016), we looked separately at anatomically defined bilat-

eral anterior and posterior hippocampal ROIs. We then used multivariate RSA (Kriegeskorte, Mur, & Bandettini, 2008) to compare the neural patterns of activity associated with encoding Object word pairs with Scene or Object single words. We predicted that the neural representations of Object word pairs in the anterior hippocampus would be more similar to Scene single words than Object single words, but that this would not be apparent in the posterior hippocampus. As our aim was to specifically investigate the contribution of different types of imagery to hippocampal activity, the scene and object single words

Table 9. Scene Word Pairs Compared with Object Word Pairs

Region	Peak Voxel Coordinates	t
A. Scene Word Pairs > Object Word Pairs		
Left anterior hippocampus	−22, −18, −20	5.55
Right anterior hippocampus	22, −20, −20	6.07
Right retrosplenial cortex	16, −54, 20	7.35
Left retrosplenial cortex	−10, −50, 4	7.34
Left fusiform cortex (extending to parahippocampal cortex)	−28, −38, −12	7.25
Right fusiform cortex (extending to parahippocampal cortex)	28, −26, −20	6.87
Left middle temporal cortex	−58, −6, −14	5.77
B. Object Word Pairs > Scene Word Pairs		
Left inferior temporal cortex	−42, −48, −16	7.16

$p < .001$ uncorrected for the hippocampus and $p < .05$ FWE-corrected for the rest of the brain. Brain regions within the medial temporal lobe were identified via visual inspection. For regions outside the medial temporal lobe, the AAL atlas was used (Tzourio-Mazoyer et al., 2002).

Table 10. Remembered Scene Word Pairs Compared with Remembered Object Word Pairs

Region	Peak Voxel Coordinates	t
A. Scene Word Pairs Remembered > Object Word Pairs Remembered		
Right hippocampus	24, −20, −20	5.18
Left hippocampus	−22, −20, −18	4.26
Left retrosplenial cortex	−12, −50, 4	6.74
Right fusiform cortex (extending to parahippocampal cortex)	24, −28, −18	6.49
Right retrosplenial cortex	10, −48, 6	6.46
Left fusiform cortex (extending to parahippocampal cortex)	−24, −38, −12	6.37
B. Object Word Pairs Remembered > Scene Word Pairs Remembered		
Left inferior temporal gyrus	−42, −48, −16	6.12

$p < .001$ uncorrected for the hippocampus and $p < .05$ FWE-corrected for the rest of the brain. Brain regions within the medial temporal lobe were identified via visual inspection. For regions outside the medial temporal lobe, the AAL atlas was used (Tzourio-Mazoyer et al., 2002).

were chosen as comparators because they consistently elicit either scene or object imagery respectively (see Methods). Abstract words do not elicit much visual imagery, so they were not included in the RSA analyses.

Three similarity correlations were calculated. First, the similarity between Object word pairs and themselves, which provided a baseline measure of similarity (i.e., the correlation of Object word pairs over the four runs of the scanning experiment). The two similarities of interest were the similarity between Object word pairs and Scene single words and the similarity between Object word pairs and Object single words. For the anterior hippocampus ROI, two participants showed similarity scores greater than 2 standard deviations away from the mean and were removed from further analysis, leaving a sample of 43 participants. For the posterior hippocampus ROI, again two participants (one of whom was also excluded from the anterior hippocampus analysis) showed similarity scores greater than 2 standard deviations away from the mean and were removed from further analysis, leaving a sample of 43 participants.

For the anterior hippocampus, a repeated-measures ANOVA found a significant difference between the three similarities, $F(2, 84) = 3.40$, $p = .038$, $\eta^2 = .075$. As predicted, the neural representations in the anterior hippocampus of Object word pairs were more similar to Scene single words (Figure 6A, purple bar) than to Object single words (Figure 6A, light green bar; $t(42) = 2.09$, $p = .042$,

$d_{rm} = 0.21$). In fact, representations of Object word pairs were as similar to Scene single words as to themselves (Figure 6A, orange bar; $t(42) = 0.38$, $p = .71$). Object word pairs were significantly less similar to Object single words than to themselves ($t(42) = 2.54$, $p = .015$, $d_{rm} = 0.23$). Of note, these results cannot be explained by subsequent memory performance because Scene single words and Object single words were remembered equally well ($t(42) = 0.68$, $p = .50$).

For the posterior hippocampus, a repeated-measures ANOVA also found a significant difference between the three similarities, $F(2, 84) = 4.83$, $p = .010$, $\eta^2 = .10$. However, in contrast to the anterior hippocampus, the neural representations in the posterior hippocampus of Object word pairs were more similar to themselves (Figure 6B, orange bar) than either Scene single words (Figure 6B, purple bar; $t(42) = 2.60$, $p = .013$, $d_{rm} = 0.32$) or Object single words (Figure 6B, light green bar; $t(42) = 2.33$, $p = .025$, $d_{rm} = 0.26$). Moreover, there was no difference between the representations of Scene and Object single words ($t(42) = -0.71$, $p = .48$). As before, these results cannot be explained by subsequent memory performance because Scene single words and Object single words were remembered equally well ($t(42) = 0.74$, $p = .46$).

Overall, these multivariate results show that, within the anterior hippocampus, Object word pairs were represented in a similar manner to Scene single words, but not Object single words. On the other hand, within the posterior hippocampus, Object word pairs were only similar to themselves. This provides further support for our hypothesis that Object word pairs evoke anterior
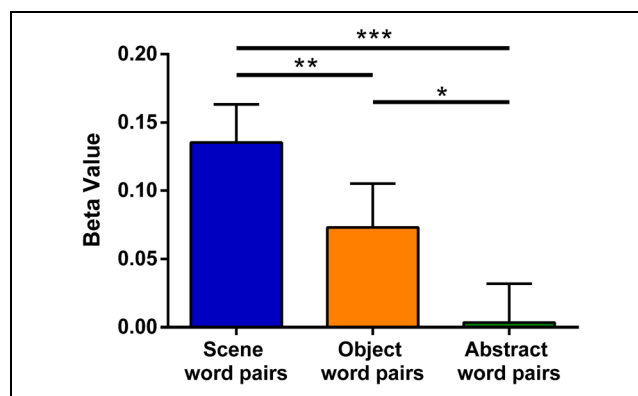

Figure 5. Comparison of each word pair condition with a fixation cross baseline. Mean beta values extracted from a bilateral anatomical mask of the anterior hippocampus for each of the word pair conditions compared with the central fixation cross baseline. Error bars are 1 *SEM*. A repeated-measures ANOVA showed significant differences between the conditions, $F(1.69, 74.51) = 16.06$, $p < .001$, $\eta^2 = .27$. Follow-up paired *t* tests revealed significant differences between Scene word pairs versus Abstract word pairs, $t(44) = 6.46$, $p < .001$, $d_{rm} = 0.70$; Scene word pairs versus Object word pairs, $t(44) = 2.97$, $p = .005$, $d_{rm} = 0.30$; Object word pairs versus Abstract word pairs, $t(44) = 2.51$, $p = .016$, $d_{rm} = 0.34$. $*p < .05$, $**p < .01$, $***p < .001$.
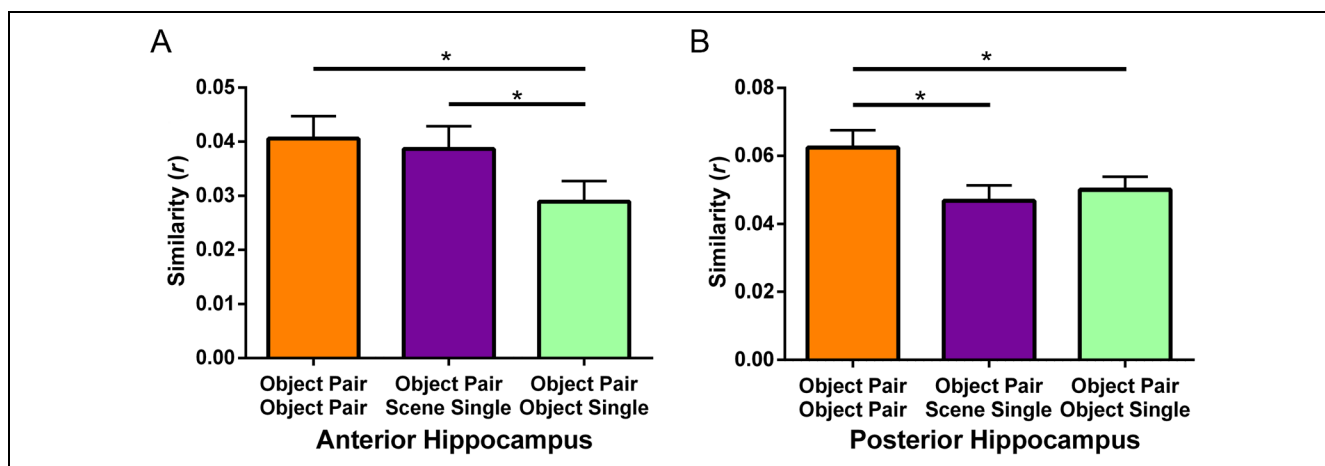
Figure 6. The neural similarity of Object word pairs, Scene single words, and Object single words separately for the anterior and posterior hippocampus. (A) Anterior hippocampus. (B) Posterior hippocampus. Object Pair Object Pair = the similarity between Object word pairs between runs; Object Pair Scene Single = the similarity between Object word pairs and Scene single words; Object Pair Object Single = the similarity between Object word pairs and Object single words. Error bars represent 1 *SEM* adjusted for repeated-measures (Morey, 2008). *$p < .05$.

(but not posterior) hippocampal activity when scene imagery is involved.

VVIQ and the Use of Imagery

As well as examining participants in one large group, as above, we also divided them into three groups based on whether they reported high, mid, or low imagery ability on the VVIQ. We found no differences in memory performance among the groups on the word pair tasks ($F < 0.4$ for all contrasts). Similarly, fMRI univariate analyses involving the word pair conditions revealed no differences in hippocampal activity. Voxel-based morphology (Ashburner, 2009; Mechelli, Price, Friston, & Ashburner, 2005; Ashburner & Friston, 2000) showed no structural differences between the groups anywhere in the brain, including in the hippocampus.

Interestingly, however, the imagery groups did differ in one specific way—their strategy for processing the Object word pairs. Although strategy use was similar across the imagery groups for the other word conditions, for the Object word pairs, twice as many participants indicated using a scene imagery strategy in the high imagery group ($n = 12/15$, 80%) than in the mid or low imagery groups ($n = 5/15$, 33% and 6/15, 40%, respectively). Comparison of scene strategy use compared with other strategy use across the imagery groups revealed this to be a significant difference, $\chi^2(2) = 7.65, p = .022$.

Given this clear difference in scene imagery use specifically for the Object word pairs, we performed the anterior and posterior hippocampus RSA analyses again for the three imagery participant groups. We hypothesized that, in the anterior hippocampus, the high imagery group would represent Object word pairs in a similar manner to Scene single words (as with our whole-group analyses), whereas this would not be the case in the mid

or low imagery groups. For the posterior hippocampus, on the other hand, we expected no differences between the imagery groups. Participants with similarity values greater than 2 standard deviations away from the mean were again excluded. For the anterior hippocampus ROI analyses, this resulted in one participant being removed from each group. For the posterior hippocampus two participants were excluded (both different participants to those excluded from the anterior hippocampus analyses), one from the mid imagery group and one from the low imagery group. Importantly, the pattern of scene imagery strategy remained the same even after the removal of these few participants (anterior hippocampus: high imagery group, $n = 11/14$; mid imagery group, $n = 5/14$; low imagery group, $n = 5/14$; $\chi^2(2) = 6.86, p = .032$; posterior hippocampus: high imagery group, $n = 12/15$; mid imagery group, $n = 4/14$; low imagery group, $n = 5/14$; $\chi^2(2) = 9.10, p = .011$).

As predicted, in the anterior hippocampus for the high imagery group, Object word pairs were more similar to Scene single words than Object single words (Figure 7A; $t(13) = 4.63, p < .001, d = 0.78$). This was not the case for the mid or low imagery groups ($t(13) = 0.472, p = .65$; $t(13) = 0.20, p = .85$, respectively). Of note, the interaction between the imagery groups was significant (Figure 7B; $F(2, 39) = 3.53, p = .039, \eta^2 = 0.15$). Independent samples t tests showed that the difference between the similarities was greater in the high imagery group than in the mid and low imagery groups ($t(26) = 2.09, p = .046, d = 0.79$ and $t(26) = 2.72, p = .011, d = 1.03$, respectively). As before, these differences cannot be explained by subsequent memory performance because all three groups showed no differences between the Scene single and Object single words (high imagery group: $t(13) = 0.35, p = .74$; mid imagery group: $t(13) = 0.40, p = .69$; low imagery group: $t(13) = 1.18, p = .26$).
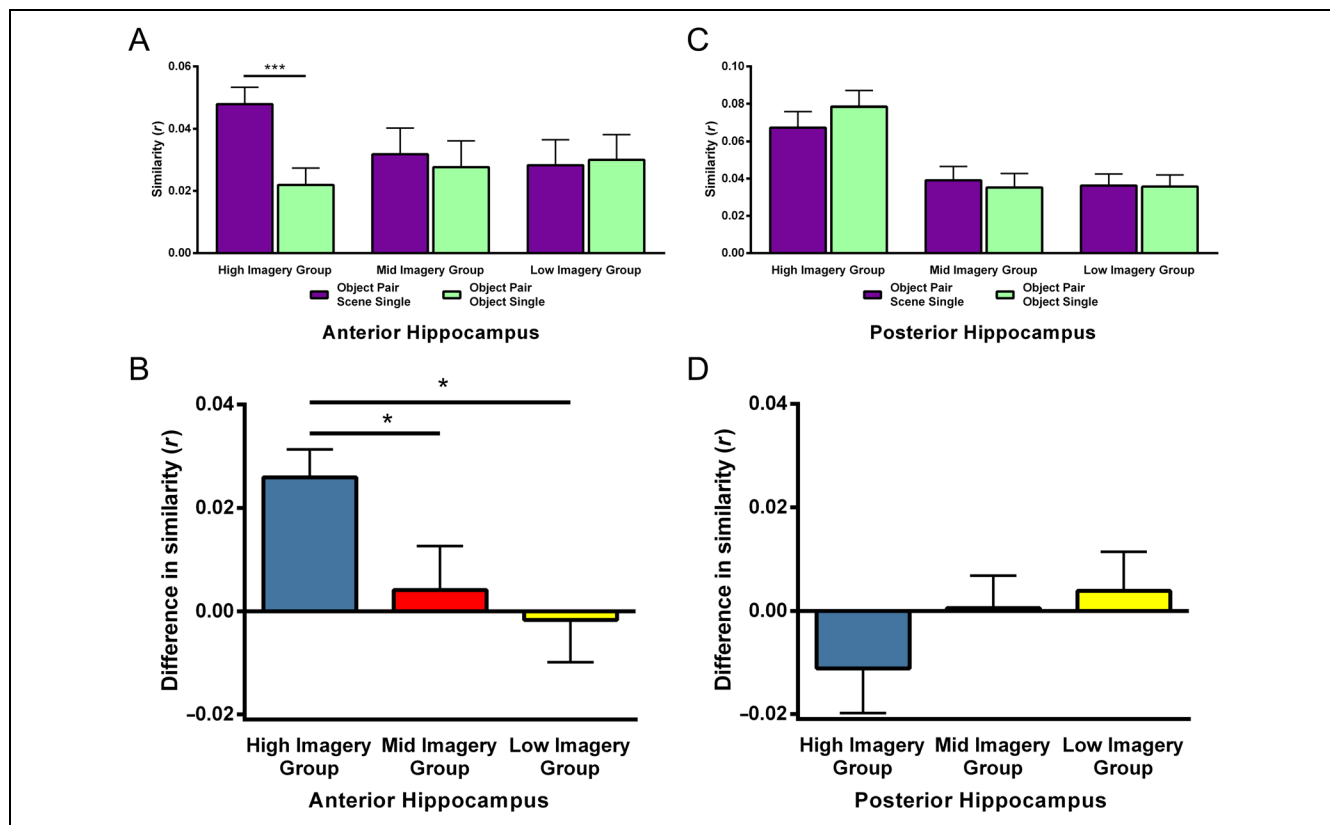
Figure 7. RSA comparisons of the three imagery groups separately for the anterior and posterior hippocampus. (A) The neural similarity of Object word pairs, Scene single words, and Object single words in the anterior hippocampus when split by self-reported imagery use. Object Pair Scene Single = the similarity between Object word pairs and Scene single words; Object Pair Object Single = the similarity between Object word pairs and Object single words. (B) The difference in similarity between Object word pairs and Scene single words compared with Object words pairs and Object single words in the imagery groups in the anterior hippocampus. (C) The neural similarity of Object word pairs, Scene single words, and Object single words in the posterior hippocampus when split by self-reported imagery use. Object Pair Scene Single = the similarity between Object word pairs and Scene single words; Object Pair Object Single = the similarity between Object word pairs and Object single words. (D) The difference in similarity between Object word pairs and Scene single words compared with Object words pairs and Object single words in the imagery groups in the posterior hippocampus. Error bars represent 1 *SEM*. *$p < .05$, ***$p < .001$.

For the posterior hippocampus, on the other hand, there were no differences in similarities in any of the imagery groups (Figure 7C; high imagery group: $t(14) = -1.29$, $p = .22$; mid imagery group: $t(13) = 0.50$, $p = .63$; low imagery group: $t(13) = 0.084$, $p = .94$). In line with these findings, the interaction between the imagery groups was also not significant (Figure 7D; $F(2, 40) = 1.07$, $p = .35$). As before, there were no differences in subsequent memory performance between the Scene single and Object single words, suggesting this was not influencing the activity patterns (high imagery group: $t(14) = 0.40$, $p = .69$; mid imagery group: $t(13) = -0.06$, $p = .95$; low imagery group: $t(13) = 1.25$, $p = .24$).

In summary, the neural patterns in anterior hippocampus for Object word pairs showed greater similarity with the Scene single words in the high imagery group, whereas for the mid and low imagery groups, this was not the case. On the other hand, we saw no differences in any of the imagery groups in the posterior hippocampus. This provides further evidence linking the anterior hippocampus with the processing of Object word pairs through scene imagery.

DISCUSSION

The aim of this study was to understand the role of the hippocampus in processing VPA. There were five findings. First, we observed greater anterior (but not posterior) hippocampal activity for high imagery (concrete) word pairs than very low imagery (abstract) word pairs, highlighting the influence of visual imagery. Second, very low imagery abstract word pairs compared with very low imagery abstract single words revealed no differences in hippocampal engagement, despite the former involving binding, adding further support for the significance of visual imagery. Third, increased anterior (but not posterior) hippocampal engagement was apparent for Scene word pairs more than Object word pairs, implicating specifically scene imagery. Fourth, for Object word pairs, fMRI activity patterns in the anterior (but not posterior) hippocampus were more similar to those for scene imagery than object imagery, further underlining the propensity of the anterior hippocampus to respond to scene imagery. Finally, our examination of high, mid, and low imagery users found that the only difference between them was the

use of scene imagery for encoding Object word pairs by high imagers, which in turn was linked to scene-related activity patterns in the anterior (but not posterior) hippocampus. Overall, our results provide evidence that anterior hippocampal engagement during VPA seems to be closely related to the use of scene imagery, even for Object word pairs.

Previous findings have hinted that visual imagery might be relevant in the hippocampal processing of verbal material such as VPA. Work in patients with right temporal lobectomies, which included removal of some hippocampal tissue, suggested that, although memory for high imagery word pairs was impaired, memory for low imagery word pairs was preserved (Jones-Gotman & Milner, 1978). Furthermore, instructing these patients to use visual imagery strategies impaired both high and low imagery word pair performance (Jones-Gotman, 1979). More recently, detailed examination of the language use of patients with bilateral hippocampal damage showed that the patients used fewer high imagery words when producing verbal narratives compared with both healthy controls and patients with damage elsewhere in the brain (Hilverman et al., 2017), supporting a link between the hippocampus and word imageability. In addition, higher than expected word pair performance has been found in amnesic patients for highly semantically related word pairs in comparison to unrelated word pairs of the kind that are usually employed in VPA tasks (Shimamura & Squire, 1984; Winocur & Weiskrantz, 1976). This suggests that when alternate strategies can be used to remember word pairs (i.e., using their semantic relationship rather than constructing scene imagery), amnesic patients do not show the typical VPA impairment. We are, however, unaware of any study that has examined VPA in patients with selective bilateral hippocampal damage where high and low imagery word pairs were directly compared (Clark & Maguire, 2016).

fMRI findings also support a possible distinction in hippocampal engagement between high and low imagery word pairs. Caplan and Madan (2016) investigated the role of the hippocampus in boosting memory performance for high imagery word pairs, concluding that imageability increased hippocampal activity. However, greater hippocampal activity for high over low imagery word pairs was only observed at a lenient whole-brain threshold ($p <$ .01 uncorrected, cluster size ≥ 5), possibly because their low imagery words (e.g., muck, fright) retained quite a degree of imageability. Furthermore, they did not examine the influence of different types of visual imagery on hippocampal engagement.

We did not find hippocampal engagement for the low imagery Abstract word pairs compared with Abstract single words, even when using ROI analyses and just the remembered stimuli. We acknowledge that null results can be difficult to interpret and that an absence of evidence is not evidence of absence. However, even our lenient uncorrected ROI analyses found no evidence of increased hippocampal activity. This is in clear contrast to the finding of increased hippocampal activity for the high imagery word pairs over the very low imagery Abstract word pairs at the whole-brain level. The most parsimonious interpretation is, therefore, that Abstract word pairs may be processed differently to the high imagery word pairs, in particular in terms of hippocampal engagement.

By contrast, activity associated with the Abstract word pairs was evident outside the hippocampus, where regions that included the left middle temporal cortex, the left temporal pole, and the left inferior frontal gyrus were engaged. These findings are in line with other fMRI studies that examined the representations of abstract words and concepts in the human brain (Wang et al., 2017; Wang, Conder, Blitzer, & Shinkareva, 2010; Binder, Westbury, McKiernan, Possing, & Medler, 2005). Our results, therefore, align with the notion of different brain systems for processing concrete (high imagery) and abstract (low imagery) concepts and stimuli.

Our different word types were extremely well matched across a wide range of features, with the abstract words being verified as eliciting very little imagery and the scene and object words as reliably eliciting the relevant type of imagery. Using these stimuli, we showed that hippocampal involvement in VPA is not linked to visual imagery in general but seems to be specifically related to scene imagery, even when each word in a pair denoted an object. This supports a prediction made by Maguire and Mullally (2013; see also Clark & Maguire, 2016), who noted that a scene allows us to collate a lot of information in a quick, coherent, and efficient manner. Consequently, they proposed that people may automatically use scene imagery during the processing of high imagery verbal material. For instance, we might visualize the scene within which a story is unfolding or place the objects described in word pairs in a simple scene together.

If verbal tasks can provoke the use of imagery-based strategies and if these strategies involve scenes, then patients with hippocampal amnesia would be expected to perform poorly on VPA tasks involving high imagery concrete words because they are known to have difficulty with constructing scenes in their imagination (e.g., Kurczek et al., 2015; Mullally, Intraub, & Maguire, 2012; Race et al., 2011; Andelman, Hoofien, Goldberg, Aizenstein, & Neufeld, 2010; Hassabis, Kumaran, Vann, et al., 2007). This impairment, which was not apparent for single objects, prompted the proposal of the scene construction theory, which holds that scene imagery constructed by the hippocampus is a vital component of memory and other functions (Maguire & Mullally, 2013; Hassabis & Maguire, 2007). Findings over the last decade have since linked scenes to the hippocampus in relation to autobiographical memory (Hassabis, Kumaran, & Maguire, 2007; Hassabis & Maguire, 2007) but also widely across cognition, including perception (McCormick et al., 2017; Mullally et al., 2012; Graham et al., 2010), future thinking (Irish, Hodges, & Piguet, 2013; Schacter et al., 2012; Hassabis, Kumaran, Vann, et al., 2007), spatial navigation (Clark & Maguire,

2016; Maguire, Nannery, & Spiers, 2006), and decision-making (McCormick, Rosenthal, Miller, & Maguire, 2016; Mullally & Maguire, 2014). However, as the current study was only designed to examine the role of the hippocampus in the VPA task, we do not speculate further here as to whether or not scene construction is the primary mechanism at play within the hippocampus. For more on this issue, we refer the reader to broader theoretical discussions of the scene construction theory (McCormick, Ciaramelli, De Luca, & Maguire, 2018; Dalton & Maguire, 2017; Clark & Maguire, 2016; Maguire, Intraub, & Mullally, 2016) and alternative accounts of hippocampal function (Moscovitch et al., 2016; Sheldon & Levine, 2016; Eichenbaum & Cohen, 2014; Schacter et al., 2012).

Our hippocampal findings were located in the anterior portion of the hippocampus. Anterior and posterior functional differentiation is acknowledged as a feature of the hippocampus, although the exact roles played by each portion are not widely agreed (Ritchey, Montchal, Yonelinas, & Ranganath, 2015; Strange, Witter, Lein, & Moser, 2014; Poppenk, Evensmoen, Moscovitch, & Nadel, 2013; Fanselow & Dong, 2010; Moser & Moser, 1998). Of note, the medial portion of the anterior hippocampus contains the presubiculum and parasubiculum hippocampal subfields. These areas have been highlighted as being consistently implicated in scene processing (reviewed in Zeidman & Maguire, 2016) and were recently proposed to be neuroanatomically determined to process scenes (Dalton & Maguire, 2017). The current results seem to accord with these findings, although higher-resolution studies are required to determine the specific subfields involved.

An important point to consider is whether our results can be explained by the effectiveness of encoding, as measured in a subsequent memory test. It is certainly true that people tend to recall fewer abstract than concrete words in behavioral studies of memory (Paivio, Walsh, & Bons, 1994; Jones, 1974; Paivio, 1969). We tested memory for both single words and paired words. Memory performance for Scene, Object, and Abstract words was comparable when tested singly. Memory for the word pairs was significantly lower for the low imagery Abstract word pairs compared with the Scene word pairs and Object word pairs. Nevertheless, performance for all conditions was above chance, which was impressive given the large number of stimuli to be encoded with only one exposure. Increased hippocampal activity was apparent for both Scene word pairs and Object words pairs compared with the Abstract word pairs when all stimuli or only the subsequently remembered stimuli were analyzed. Furthermore, although Object word pairs were remembered better than Scene word pairs, hippocampal activity was nevertheless greater for the Scene word pairs. This shows that our results cannot be explained by encoding success. It is also worth considering why such a gradient in memory performance was observed within the word pairs, with Object word pairs being remembered better than Scene

word pairs, which were remembered better than Abstract word pairs. One possibility may be that memory performance benefitted from the extent to which the two words could be combined into some kind of relationship. This is arguably easier for two objects than two scenes, both of which are easier than for two abstract concepts.

Although differences between the performance of amnesic patients and healthy participants on VPA tasks are typically observed during cued recall, in the current study we used recognition memory tests postscanning to assess the success of encoding. This is because testing cued recall for 135 word pairs that were each seen only once is simply too difficult even for healthy participants. For example, learning just 14 (high imagery concrete) word pairs on the WMS-IV VPA task is performed over four learning trials. We did, however, ensure that the associative recognition memory test was challenging by constructing the lure word pairs from the single words that were presented to the participants during scanning. Thus, all words were previously seen by participants, but not all were previously seen in pairs.

Moreover, we believe that the use of a recognition memory test instead of cued recall had little impact on the patterns of brain activity we observed because brain activity was assessed during the initial presentation of the word pairs and not during memory retrieval. As participants were not told exactly how their memory would be tested after the learning phase, it might be expected that participants engaged in the most effortful encoding that they could. That the involvement of the hippocampus was identified when using all the trials in the fMRI analysis or just the subsequently remembered stimuli also points to the use of imagery at the time of stimuli presentation as being of most relevance rather than encoding success.

There is a wealth of research linking the hippocampus with associative binding (e.g., Palombo, Hayes, Peterson, Keane, & Verfaellie, 2018; Roberts et al., 2018; Rangel et al., 2016; Schwarb et al., 2015; Eichenbaum & Cohen, 2014; Addis, Cheng, Roberts, & Schacter, 2011; Konkel & Cohen, 2009; Davachi, 2006). We do not deny this is the case but suggest that our results provoke a reconsideration of the underlying reason for apparent associative effects. We found that the creation of associations between nonimageable Abstract word pairs did not elicit an increase in hippocampal activity compared with Abstract single words, even when only subsequently remembered stimuli were considered. If binding per se was the reason for hippocampal involvement in our study, then this contrast should have revealed it. We suggest instead that the anterior hippocampus engages in associative binding specifically to create scene imagery and that this relationship with scenes has been underestimated or ignored in VPA and other associative tasks despite potentially having a significant influence on hippocampal engagement.

Our participants were self-declared low, mid, or high imagery users as measured by the VVIQ. They differed

only in the degree of scene imagery usage, in particular during the processing of Object word pairs, with high imagers showing the greatest amount. Given that scene imagery has been implicated in functions across cognition, it might be predicted that those who are able to use scene imagery well might have more successful recall of autobiographical memories and better spatial navigation. Individual differences studies are clearly required to investigate this important issue in depth, as currently there is a dearth of such work (a point also noted by Palombo, Sheldon, & Levine, 2018). In this study, increased use of scene imagery by the high imagery group did not convey a memory advantage for the Object word pairs. However, in the real world, with more complex memoranda like autobiographical memories, we predict that scene imagery would promote better memory.

In conclusion, we showed a strong link between the anterior hippocampus and processing words in a VPA task mediated through scene imagery. This offers a way to reconcile hippocampal theories that have a visuo-spatial bias with the processing and subsequent memory of verbal material. Moreover, we speculate that this could hint at a verbal system in humans piggy-backing on top of an evolutionarily older visual (scene) mechanism (see also Corballis, 2017). We believe it is likely that other common verbal tests, such as story recall and list learning, which are typically highly imageable, may similarly engage scene imagery and the anterior hippocampus. Greater use of low imagery abstract verbal material would seem to be prudent in future verbal memory studies. Indeed, an obvious prediction arising from our results is that patients with selective bilateral hippocampal damage would be better at recalling abstract compared with imageable word pairs, provided care is taken to match the stimuli precisely. Our data do not speak to the issue of whether or not scene construction is the primary mechanism at play within the hippocampus, as our main interest was in examining VPA, a task closely aligned with the hippocampus. What our results show, and we believe in a compelling fashion, is that anterior hippocampal engagement during VPA seems to be best explained by the use of scene imagery.

Appendix A. List of Scene Words Used in the Study

Airfield	Coast	Island	Ravine
Airport	Cockpit	Jungle	Reservoir
Aisle	Coliseum	Kitchen	Restaurant
Alley	College	Laboratory	River
Apartment	Constellation	Lagoon	Road
Aquarium	Corridor	Lake	Rodeo
Arcade	Cottage	Landscape	Rooftop
Arena	Countryside	Lane	Salon
Attic	Courtroom	Lawn	Sandstorm
Auditorium	Courtyard	Library	Sauna
Avalanche	Cove	Lightning	School
Avenue	Crater	Loft	Sea
Backyard	Creek	Manor	Seaside
Bakery	Crowd	Market	Sewer
Ballroom	Crypt	Mausoleum	Shipwreck
Bank	Cubicle	Maze	Shipyard
Banquet	Dawn	Meadow	Shop
Barbecue	Depot	Monastery	Sky
Barnyard	Dock	Morgue	Snowstorm
Basement	Dormitory	Motel	Stadium
Bathroom	Drawbridge	Mountain	Stage
Battlefield	Driveway	Museum	Station
Bay	Dungeon	Newsroom	Stockroom
Beach	Eclipse	Nightclub	Storm
Bedroom	Explosion	Nursery	Stream
Blizzard	Factory	Oasis	Street
Boardroom	Farm	Observatory	Studio
Bog	Festival	Ocean	Sunrise
Bookshop	Fireworks	Office	Sunset
Boutique	Fjord	Orchard	Supermarket
Brewery	Flood	Orchestra	Swamp
Brook	Forest	Palace	Tavern
Buffet	Fortress	Parade	Temple
Cabin	Foyer	Park	Terrace
Cafeteria	Gallery	Passageway	Terrain
Campsite	Garage	Pasture	Theatre
Canyon	Garden	Patio	Tornado
Carnival	Glacier	Peninsula	Tournament
Casino	Gorge	Pharmacy	Tower
Castle	Graveyard	Picnic	Town
Cathedral	Gym	Planetarium	Trail
Catwalk	Gymnasium	Plantation	Tunnel
Cave	Hairdresser	Plateau	Twilight
Cellar	Hall	Playground	University
Cemetery	Harbor	Playroom	Valley
Chapel	Heliport	Plaza	Villa
Chateau	Herd	Pond	Village
Choir	Highway	Port	Vineyard
Church	Hill	Prairie	Volcano
Cinema	Horizon	Pub	Warehouse
Circus	Hospital	Quarry	Waterfall
City	Hotel	Racetrack	Waterfront
Classroom	Hurricane	Railroad	Zoo
Cliff	Infirmary	Railway	
Clinic	Inn	Ranch	

Appendix B. List of Object Words Used in the Study

Ambulance	Cod	Lifeboat	Rose
Amulet	Coffee	Limousine	Ruby
Apron	Comet	Lobster	Sapphire
Asterisk	Compass	Locket	Satellite
Badger	Computer	Lotus	Saucer
Bandage	Conditioner	Machine	Scale
Banjo	Container	Magazine	Scalpel
Barrel	Cooker	Marmalade	Scone
Bayonet	Crepe	Medal	Screwdriver
Beer	Crocodile	Medallion	Seat
Bench	Crown	Menu	Shampoo
Biscuit	Crucifix	Microphone	Shell
Blazer	Cushion	Microwave	Shrapnel
Blindfold	Cylinder	Minibus	Shutter
Blossom	Desk	Monkey	Signature
Book	Diamond	Mosquito	Skirt
Boomerang	Dice	Motorbike	Skull
Bouquet	Dictionary	Muffin	Sleigh
Bowl	Dough	Nest	Smoothie
Brick	Doughnut	Newspaper	Snorkel
Broccoli	Dragon	Noose	Socket
Brooch	Drainpipe	Nut	Soup
Bubble	Driftwood	Olive	Spade
Buckle	Drill	Ostrich	Staple
Bullet	Dynamite	Package	Stretcher
Buoy	Embroidery	Packet	Submarine
Butter	Emerald	Page	Sunflower
Buttercup	Envelope	Painting	Swing
Button	Espresso	Pallet	Tabloid
Cabinet	Fairy	Pamphlet	Tangerine
Calf	Fan	Panda	Tapestry
Camera	Flask	Parcel	Telescope
Canoe	Frame	Parchment	Thread
Cap	Generator	Pasta	Ticket
Card	Guitar	Pendulum	Topaz
Cardinal	Hammer	Pepper	Tortoise
Carp	Hamper	Peppermint	Treasure
Carriage	Handle	Photograph	Tripod
Catalogue	Harness	Picture	Trolley
Catapult	Harpoon	Pint	Tulip
Cent	Headphones	Potato	Turnip
Certificate	Heart	Projector	Typewriter
Chain	Helicopter	Prune	Van
Champagne	Honey	Pterodactyl	Vase
Chariot	Honeycomb	Pudding	Vodka
Cheese	Hook	Pump	Wand
Cheetah	Horseradish	Pussycat	Weasel
Cherry	Jigsaw	Rabbit	Wedge
Chestnut	Kangaroo	Racehorse	Wheelbarrow
Chisel	Kayak	Radiator	Whip
Cigarette	Ladle	Raspberry	Whistle
Cloak	Leech	Recorder	Window
Clock	Lemon	Reptile	Witch
Coaster	Leopard	Rifle	
Cocktail	Letter	Rocket	

Appendix C. List of Abstract Words Used in the Study

Abstinence	Ego	Logic	Relevance
Absurd	Empathy	Luck	Remorse
Accord	Envy	Manner	Renown
Affront	Essence	Meaning	Repression
Ambition	Esteem	Memory	Reproach
Amends	Eternity	Mercy	Resolve
Annual	Ethic	Merit	Respect
Aptitude	Euphemism	Metaphor	Respite
Aspect	Existence	Method	Retrospect
Assurance	Expertise	Midst	Reverence
Attempt	Extent	Mistake	Rhetoric
Attitude	Extreme	Mode	Riddance
Avail	Fairness	Moment	Risk
Awareness	Fate	Mood	Role
Basis	Fault	Moral	Rumour
Behavior	Feel	Morale	Sanctity
Belief	Feeling	Motive	Sanity
Bias	Finesse	Mystique	Sarcasm
Blame	Folly	Myth	Seriousness
Bother	Foresight	Need	Skill
Care	Forgiveness	Neutral	Snub
Cause	Function	Nonsense	Splendour
Certainty	Gain	Normal	Standard
Chance	Gist	Notion	Standpoint
Clarity	Godsend	Opinion	Stoic
Closure	Grandeur	Origin	Strategy
Commitment	Grudge	Outcome	Subconscious
Concept	Guess	Oversight	Succession
Concern	Guilt	Paradigm	Taboo
Conduct	Hint	Paradox	Tact
Confidence	Honesty	Pardon	Tendency
Conjecture	Honour	Patience	Theory
Conscience	Idea	Pause	Think
Contention	Ideal	Penance	Thrift
Context	Idealist	Piety	Tradition
Courtesy	Importance	Pity	Transition
Creed	Infinity	Plea	Trust
Cunning	Insight	Precaution	Truth
Debut	Intellect	Prestige	Try
Decision	Intent	Principles	Uncertainty
Decorum	Intention	Prophecy	Unknown
Default	Interest	Prudence	Utmost
Desire	Intuition	Psyche	Validity
Despair	Involvement	Pun	Value
Destiny	Irony	Purity	Version
Difference	Issue	Purpose	Virtue
Dignity	Judgment	Quality	Way
Dilemma	Karma	Rarity	Whim
Discretion	Kind	Readiness	Willpower
Distinction	Knack	Realism	Wisdom
Distrust	Lack	Reason	Wish
Doubt	Leeway	Reasoning	Woe
Dread	Legacy	Reckoning	Zeal
Duty	Leniency	Reform	
Effect	Likelihood	Regard	

REFERENCES

Addis, D. R., Cheng, T., Roberts, R. P., & Schacter, D. L. (2011). Hippocampal contributions to the episodic simulation of specific and general future events. *Hippocampus, 21,* 1045–1052.

Andelman, F., Hoofien, D., Goldberg, I., Aizenstein, O., & Neufeld, M. Y. (2010). Bilateral hippocampal lesion and a selective impairment of the ability for mental time travel. *Neurocase, 16,* 426–435.

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron, 65,* 550–562.

Ashburner, J. (2009). Computational anatomy with the SPM software. *Magnetic Resonance Imaging, 27,* 1163–1174.

Ashburner, J., & Friston, K. J. (2000). Voxel based morphometry —The methods. *Neuroimage, 11,* 805–821.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Bellmund, J. L. S., Deuker, L., Navarro Schröder, T., & Doeller, C. F. (2016). Grid-cell representations in mental simulation. *eLife, 5,* e17089.

Binder, J. R., Westbury, C., McKiernan, K., Possing, E., & Medler, D. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience, 17,* 905–917.

Brysbaert, M., Warriner, A., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46,* 904–911.

Burton, L. J., & Fogarty, G. J. (2003). The factor structure of visual imagery and spatial abilities. *Intelligence, 31,* 289–318.

Campos, A., & Pérez-Fabello, M. J. (2005). The Spanish version of Betts' Questionnaire upon Mental Imagery. *Psychological Reports, 96,* 51–56.

Caplan, J. B., & Madan, C. R. (2016). Word imageability enhances association-memory by increasing hippocampal engagement. *Journal of Cognitive Neuroscience, 28,* 1522–1538.

Chadwick, M. J., Jolly, A. E. J., Amos, D. P., Hassabis, D., & Spiers, H. J. (2015). A goal direction signal in the human entorhinal/subicular region. *Current Biology, 25,* 87–92.

Clark, I. A., & Maguire, E. A. (2016). Remembering preservation in hippocampal amnesia. *Annual Review of Psychology, 67,* 51–82.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia and the hippocampal system.* Cambridge, MA: MIT Press.

Corballis, M. C. (2017). Language evolution: A changing perspective. *Trends in Cognitive Sciences, 21,* 229–236.

Dalton, M. A., & Maguire, E. A. (2017). The pre/parasubiculum: A hippocampal hub for scene-based cognition? *Current Opinion in Behavioral Sciences, 17,* 34–40.

Dalton, M. A., Zeidman, P., Barry, D. N., Williams, E., & Maguire, E. A. (2017). Segmenting subregions of the human hippocampus on structural magnetic resonance image scans: An illustrated tutorial. *Brain and Neuroscience Advances, 1,* 1–36.

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

Deichmann, R., Schwarzbauer, C., & Turner, R. (2004). Optimisation of the 3D MDEFT sequence for anatomical brain imaging: Technical implications at 1.5 and 3 T. *Neuroimage, 21,* 757–767.

Deuker, L., Bellmund, J. L. S., Navarro Schröder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *eLife, 5,* e16534.

Eichenbaum, H., & Cohen, N. J. (2014). Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron, 83,* 764–770.

Fanselow, M. S., & Dong, H.-W. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron, 65,* 7–19.

Giovanello, K. S., Verfaellie, M., & Keane, M. M. (2003). Disproportionate deficit in associative recognition relative to item recognition in global amnesia. *Cognitive, Affective, & Behavioral Neuroscience, 3,* 186–194.

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.

Graham, K. S., Barense, M. D., & Lee, A. C. H. (2010). Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia, 48,* 831–853.

Gur, R. C., & Hilgard, E. R. (1975). Visual imagery and the discrimination of differences between altered pictures simultaneously and successively presented. *British Journal of Psychology, 66,* 341–345.

Haist, F., Shimamura, A. P., & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 691–702.

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience, 27,* 14365–14374.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences, U.S.A., 104,* 1726–1731.

Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences, 11,* 299–306.

Henson, R. N., Horner, A. J., Greve, A., Cooper, E., Gregori, M., Simons, J. S., et al. (2017). No effect of hippocampal lesions on stimulus-response bindings. *Neuropsychologia, 103,* 106–114.

Hilverman, C., Cook, S. W., & Duff, M. C. (2017). The influence of the hippocampus and declarative memory on word use: Patients with amnesia use less imageable words. *Neuropsychologia, 106,* 179–186.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of Computer Assisted Tomography, 22,* 324–333.

Hsieh, L.-T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron, 81,* 1165–1178.

Hsieh, L.-T., & Ranganath, C. (2015). Cortical and subcortical contributions to sequence retrieval: Schematic coding of temporal context in the neocortical recollection network. *Neuroimage, 121,* 78–90.

Irish, M., Hodges, J. R., & Piguet, O. (2013). Episodic future thinking is impaired in the behavioural variant of frontotemporal dementia. *Cortex, 49,* 2377–2388.

Jones, M. K. (1974). Imagery as a mnemonic aid after left temporal lobectomy: Contrast between material-specific and generalized memory disorders. *Neuropsychologia, 12,* 21–30.

Jones-Gotman, M. (1979). Incidental learning of image-mediated or pronounced words after right temporal lobectomy. *Cortex, 15,* 187–197.

Jones-Gotman, M., & Milner, B. (1978). Right temporal-lobe contribution to image-mediated verbal learning. *Neuropsychologia, 16,* 61–71.

Kim, M., Jeffery, K. J., & Maguire, E. A. (2017). Multivoxel pattern analysis reveals 3D place information in the human hippocampus. *Journal of Neuroscience, 37,* 4270–4279.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273,* 1399–1402.

Konkel, A., & Cohen, N. J. (2009). Relational memory and the hippocampus: Representations and methods. *Frontiers in Neuroscience, 3,* 166–174.

Kosslyn, S. M., Brunn, J., Cave, K. R., & Wallach, R. W. (1984). Individual differences in mental-imagery ability: A computational analysis. *Cognition, 18,* 195–243.

Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General, 140,* 14–34.

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience, 2,* 4.

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., et al. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron, 60,* 1126–1141.

Kumaran, D., Hassabis, D., Spiers, H. J., Vann, S. D., Vargha-Khadem, F., & Maguire, E. A. (2007). Impaired spatial and non-spatial configural learning in patients with hippocampal pathology. *Neuropsychologia, 45,* 2699–2711.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44,* 978–990.

Kurczek, J., Wechsler, E., Ahuja, S., Jensen, U., Cohen, N. J., Tranel, D., et al. (2015). Differential contributions of hippocampus and medial prefrontal cortex to self-projection and self-referential processing. *Neuropsychologia, 73,* 116–126.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t* tests and ANOVAs. *Frontiers in Psychology, 4,* 863.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., et al. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 4398–4403.

Maguire, E. A., Intraub, H., & Mullally, S. L. (2016). Scenes, spaces, and memory traces: What does the hippocampus do? *The Neuroscientist, 22,* 432–439.

Maguire, E. A., & Mullally, S. L. (2013). The hippocampus: A manifesto for change. *Journal of Experimental Psychology: General, 142,* 1180–1189.

Maguire, E. A., Nannery, R., & Spiers, H. J. (2006). Navigation around London by a taxi driver with bilateral hippocampal lesions. *Brain, 129,* 2894–2907.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences, U.S.A., 92,* 8135–8139.

Marchette, S. A., Vass, L. K., Ryan, J., & Epstein, R. A. (2014). Anchoring the neural compass: Coding of local spatial reference frames in human medial parietal lobe. *Nature Neuroscience, 17,* 1598–1606.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology, 64,* 17–24.

Mayes, A. R., Holdstock, J. S., Isaac, C. L., Montaldi, D., Grigor, J., Gummer, A., et al. (2004). Associative recognition in a patient with selective hippocampal lesions and relatively normal item recognition. *Hippocampus, 14,* 763–784.

Mayes, A. R., Isaac, C. L., Holdstock, J. S., Hunkin, N. M., Montaldi, D., Downes, J. J., et al. (2001). Memory for single items, word pairs, and temporal order of different kinds in a patient with selective hippocampal lesions. *Cognitive Neuropsychology, 18,* 97–123.

McAvinue, L. P., & Robertson, I. H. (2007). Measuring visual imagery ability: A review. *Imagination, Cognition and Personality, 26,* 191–211.

McCormick, C., Ciaramelli, E., De Luca, F., & Maguire, E. A. (2018). Comparing and contrasting the cognitive effects of hippocampal and ventromedial prefrontal cortex damage: A review of human lesion studies. *Neuroscience, 374,* 295–318.

McCormick, C., Rosenthal, C. R., Miller, T. D., & Maguire, E. A. (2016). Hippocampal damage increases deontological responses during moral decision making. *Journal of Neuroscience, 36,* 12157–12167.

McCormick, C., Rosenthal, C. R., Miller, T. D., & Maguire, E. A. (2017). Deciding what is possible and impossible following hippocampal damage in humans. *Hippocampus, 27,* 303–314.

Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2005). Voxel based morphometry of the human brain: Methods and applications. *Current Medical Imaging Reviews, 1,* 105–113.

Milivojevic, B., Vicente-Grabovetsky, A., & Doeller, C. F. (2015). Insight reconfigures hippocampal-prefrontal memories. *Current Biology, 25,* 821–830.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4,* 61–64.

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic memory and beyond: The hippocampus and neocortex in transformation. *Annual Review of Psychology, 67,* 105–134.

Moser, M.-B., & Moser, E. I. (1998). Functional differentiation in the hippocampus. *Hippocampus, 8,* 608–619.

Mullally, S. L., Intraub, H., & Maguire, E. A. (2012). Attenuated boundary extension produces a paradoxical memory advantage in amnesic patients. *Current Biology, 22,* 261–268.

Mullally, S. L., & Maguire, E. A. (2014). Counterfactual thinking in patients with amnesia. *Hippocampus, 24,* 1261–1266.

Murray, E. A., Bussey, T. J., & Saksida, L. M. (2007). Visual perception and memory: A new view of medial temporal lobe function in primates and rodents. *Annual Review of Neuroscience, 30,* 99–122.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map.* Oxford: Clarendon Press.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76,* 241–263.

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory—When and why. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1196–1204.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76,* 1–25.

Palombo, D. J., Hayes, S. M., Peterson, K. M., Keane, M. M., & Verfaellie, M. (2018). Medial temporal lobe contributions to

episodic future thinking: Scene construction or future projection? *Cerebral Cortex, 28,* 447–458.

Palombo, D. J., Sheldon, S., & Levine, B. (2018). Individual differences in autobiographical memory. *Trends in Cognitive Sciences, 22,* 583–597.

Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences, 17,* 230–240.

Race, E., Keane, M. M., & Verfaellie, M. (2011). Medial temporal lobe damage causes deficits in episodic memory and episodic future thinking not attributable to deficits in narrative construction. *Journal of Neuroscience, 31,* 10262–10269.

Rangel, L. M., Rueckemann, J. W., Riviere, P. D., Keefe, K. R., Porter, B. S., Heimbuch, I. S., et al. (2016). Rhythmic coordination of hippocampal neurons during associative memory processing. *eLife, 5,* e09849.

Ritchey, M., Montchal, M. E., Yonelinas, A. P., & Ranganath, C. (2015). Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. *eLife, 4,* e05025.

Roberts, R. P., Schacter, D. L., & Addis, D. R. (2018). Scene construction and relational processing: Separable constructs? *Cerebral Cortex, 28,* 1729–1732.

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology, 12,* 191–200.

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: Remembering, imagining, and the brain. *Neuron, 76,* 677–694.

Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus, 26,* 3–8.

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. *Neuron, 91,* 1402–1412.

Schwarb, H., Watson, P. D., Campbell, K., Shander, C. L., Monti, J. M., Cooke, G. E., et al. (2015). Competition and cooperation among relational memory representations. *PLoS One, 10,* e0143832.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry, 20,* 11–21.

Sheehan, P. W. (1967). A shortened form of Betts Questionnaire upon Mental Imagery. *Journal of Clinical Psychology, 23,* 386–389.

Sheldon, S., & Levine, B. (2016). The role of the hippocampus in memory and mental construction. *Annals of the New York Academy of Sciences, 1369,* 76–92.

Shimamura, A. P., & Squire, L. R. (1984). Paired-associate learning and priming effects in amnesia: A neuropsychological study. *Journal of Experimental Psychology: General, 113,* 556–570.

Spiers, H. J., Maguire, E. A., & Burgess, N. (2001). Hippocampal amnesia. *Neurocase, 7,* 357–382.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99,* 195–231.

St. Jacques, P. L., Carpenter, A. C., Szpunar, K. K., & Schacter, D. L. (2018). Remembering and imagining alternative versions of the personal past. *Neuropsychologia, 110,* 170–179.

St. Jacques, P. L., Conway, M. A., Lowder, M. W., & Cabeza, R. (2010). Watching my mind unfold versus yours: An fMRI study using a novel camera technology to examine neural differences in self-projection of self versus other perspectives. *Journal of Cognitive Neuroscience, 23,* 1275–1284.

Stadthagen-Gonzalez, H., & Davis, C. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods, 38,* 598–605.

Staresina, B. P., Henson, R. N. A., Kriegeskorte, N., & Alink, A. (2012). Episodic reinstatement in the medial temporal lobe. *Journal of Neuroscience, 32,* 18150–18156.

St-Laurent, M., Moscovitch, M., & McAndrews, M. P. (2016). The retrieval of perceptual memory details depends on right hippocampal integrity and activation. *Cortex, 84,* 15–33.

Strange, B. A., Witter, M. P., Lein, E. S., & Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. *Nature Reviews Neuroscience, 15,* 655–669.

Thakral, P. P., Benoit, R. G., & Schacter, D. L. (2017). Characterizing the role of the hippocampus during episodic simulation and encoding. *Hippocampus, 27,* 1275–1284.

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.

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67,* 1176–1190.

Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., et al. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex, 24,* 1767–1777.

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, X., Wu, W., Ling, Z., Xu, Y., Fang, Y., Wang, X., et al. (2017). Organizational principles of abstract words in the human brain. *Cerebral Cortex,* 1–14. doi:10.1093/cercor/bhx283.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45,* 1191–1207.

Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology, 19,* 87–95.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV).* San Antonio, TX: NCS Pearson.

Wechsler, D. (2009). *WMS-IV: Wechsler Memory Scale-Administration and scoring manual.* London: Psychological Corporation.

Wechsler, D. (2011). *Test of Premorbid Functioning–UK version (TOPF UK).* London: Pearson Assessment.

Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3 T and 1.5 T. *Neuroimage, 33,* 493–504.

Winocur, G., & Weiskrantz, L. (1976). An investigation of paired-associate learning in amnesic patients. *Neuropsychologia, 14,* 97–110.

Zeidman, P., & Maguire, E. A. (2016). Anterior hippocampus: The anatomy of perception, imagination and episodic memory. *Nature Reviews Neuroscience, 17,* 173–182.

Zola-Morgan, S., Squire, L. R., & Amaral, D. G. (1986). Human amnesia and the medial temporal region: Enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. *Journal of Neuroscience, 6,* 2950–2967.