The Effect of Scene Context on Episodic Object Recognition: Parahippocampal Cortex Mediates Memory Encoding and Retrieval Success

Scott M. Hayes,1 Lynn Nadel,2* and Lee Ryan2

ABSTRACT: Previous research has investigated intentional retrieval of contextual information and contextual influences on object identification and word recognition, yet few studies have investigated context effects in episodic memory for objects. To address this issue, unique objects embedded in a visually rich scene or on a white background were presented to participants. At test, objects were presented either in the original scene or on a white background. A series of behavioral studies with young adults demonstrated a context shift decrement (CSD)—decreased recognition performance when context is changed between encoding and retrieval. The CSD was not attenuated by encoding or retrieval manipulations, suggesting that binding of object and context may be automatic. A final experiment explored the neural correlates of the CSD, using functional Magnetic Resonance Imaging. Parahippocampal cortex (PHC) activation (right greater than left) during incidental encoding was associated with subsequent memory of objects in the context shift condition. Greater activity in right PHC was also observed during successful recognition of objects previously presented in a scene. Finally, a subset of regions activated during scene encoding, such as bilateral PHC, was reactivated when the object was presented on a white background at retrieval. Although participants were not required to intentionally retrieve contextual information, the results suggest that PHC may reinstate visual context to mediate successful episodic memory retrieval. The CSD is attributed to automatic and obligatory binding of object and context. The results suggest that PHC is important not only for processing of scene information, but also plays a role in successful episodic memory encoding and retrieval. These findings are consistent with the view that spatial information is stored in the hippocampal complex, one of the central tenets of Multiple Trace Theory. © 2007 Wiley-Liss, Inc.

KEY WORDS: episodic memory; parahippocampal gyrus; hippocampus; context; multiple trace theory

INTRODUCTION

The term context is often used to refer to information that is present in the environment but is irrelevant or at least incidental to the cognitive task being performed. The influence of context on memory has been investigated almost exclusively in studies using verbal materials, with changes in context between learning and retrieval having a detrimental effect on memory performance (Gold and Baddeley, 1975; Eich, 1985; Smith and Guthrie, 1921 cited in Dulsky, 1935; for a review, see Smith, 1988), although results were not always consistent (Smith et al., 1978; Godden and Baddeley, 1980; Fernandez and Glenberg, 1985). The detrimental effect of context change initially appeared to be stronger for recall than recognition; however, a meta-analysis by Smith and Vela (2001) indicated similar effect sizes (~0.27) for environmental context manipulations irrespective of test type (free recall, cued recall, or recognition). The aforementioned experiments presented verbal materials in one of only two environmental or global contexts (underwater or on land), and then memory was tested either in the same or the alternate environmental context.

Other contextual manipulations have focused on more local aspects of visual context combined with verbal materials, such as text color, background color, or font. Dulskey (1935), in an elegant series of experiments, reported a decrease in memory performance when the background color of target nonsense syllables changed between study and test. Since then, many experiments have demonstrated decreased memory performance with changes between encoding and retrieval in the local verbal context (Tulving and Osler, 1968; Light and Carter-Sobell, 1970), font format and orientation (Graf and Ryan, 1990), background color (Mori and Graf, 1996), or foreground and background color (Dougal and Rotello, 1999). In a comprehensive series of experiments, Murnane and Phelps (1993, 1994, 1995) manipulated context by changing foreground (color of the word), background (color of computer screen), and the location of the word (upper left, lower right, etc.). In multiple experiments, a context shift decrement (CSD)—decreased memory for items presented in different contexts at study and test—was observed. The CSD was significantly enhanced when the words were originally studied in a visually rich context (computer-generated virtual reality scenes, such as on a chalkboard in a classroom) relative to simple visual contexts (colored font, colored background, or in various locations on the computer screen; Murnane et al., 1999).
Surprisingly, few studies have investigated the influence of context on object recognition (Park et al., 1984, 1987) investigated context effects and aging, and a handful of studies have used faces (e.g., Winograd and Rivers-Bulkeley, 1977; Smith and Vela, 1986)). A notable exception is a recent report by Hollingworth (2006), who tested object recognition in the same context or on an olive colored background after a 20 or 40-s delay, and observed a CSD of 8.5 and 5.2%, respectively. Evidence for the role of context in object memory comes primarily from research on object perception or object identification in humans, which has shown that contextual information enhances object identification (Palmer, 1975; Biederman et al., 1982; Boyce and Pollatsek, 1992; Davenport and Potter, 2004). For example, Bar and Ullman (1996) showed that the presence of a clearly identifiable object facilitated identification of an ambiguous object when the identifiable object was semantically related, as did the presentation of realistic spatial relationships between related objects. Bar (2003, 2004) proposed specific mechanisms underlying contextual facilitation of object identification, hypothesizing that low frequency visual information is quickly extracted from a visual scene and leads to activation of a “context frame.” Context frames are defined as “prototypical representations of unique contexts (for example, a library), which guide the formation of specific instantiations in episodic scenes (for example, our library)” (Bar, 2004; p 618). The context frame, in turn, serves to reduce the number of plausible objects within a given scene and activates associated object representations that facilitate object identification. Importantly, the relationship between an object and context frame is bidirectional. Context frames activate object representations, and objects activate context frames. For instance, some objects (a basketball) are highly associated with a specific context (a gymnasium), whereas other objects (a dog) are not.

These experiments focused on pre-existing, semantic relationships between objects and their associated contexts, providing strong evidence that contextual information enhances the identification of a class or type of object. However, the influence of visual context on memory for a specific, episodically presented object remains to be determined. In one relevant series of experiments, Gaffan (1994) documented an interaction between object and spatial contextual information in episodic-like memory in non-human primates. Normal monkeys learned “object-in-place” discrimination faster than object discrimination or place discrimination alone. Fornix-lesioned monkeys were more impaired on “object-in-place” trials than on either place discrimination or object discrimination trials alone. Gaffan suggested that memory for object and context information conjoined differs from memory for object or context information alone, but that the mechanisms underlying all these forms of memory reside within medial temporal lobe structures, consistent with rodent studies suggesting neuroanatomical and behavioral dissociations in memory for object and contextual information (O’Keefe and Nadel, 1978; Nadel and Willner, 1980; Eacott and Norman, 2004; Norman and Eacott, 2005).

Recent neuroimaging studies in humans have implicated parahippocampal cortex (PHC) as a region critical for scene processing (Epstein and Kanwisher, 1998; Epstein, 2005). Using functional Magnetic Resonance Imaging (fMRI), Bar and Aminoff (2003) observed increased activation in PHC and retrosplenial cortex when participants viewed familiar objects with strong spatial contextual associations. Similarly, Epstein et al. (1999; Experiment 1) presented pictures of landmarks (on a white background) from M.I.T. and Tufts campuses to students from each respective university. Familiar landmarks (e.g., M.I.T. landmarks to an M.I.T student) elicited greater PHC activation than unfamiliar landmarks (e.g., Tufts landmarks to an M.I.T. student). In both studies, increased PHC activation was attributed to processing contextual (scene) information associated with the familiar objects. Epstein and Higgins (2006) recently found differential PHC and retrosplenial cortex activation based on the “type” of scene to be processed, e.g., a general place (garage), a familiar specific place (Yankee Stadium), or a situational category (party); however, a condition to assess the role of PHC in relation to a specific episode was not included in their study.

Hayes et al. (2004) demonstrated that PHC plays a role in intentional retrieval of spatial contextual information. They observed preferential activation in right PHC during episodic retrieval of spatial location information—more so than during the intentional retrieval of either object or temporal order information. The results suggested that PHC activation was not driven simply by the presentation of complex scenes, but by the specific requirements of the episodic memory task, as PHC responded preferentially only when subjects were required to remember the combination of scene elements that had been presented at study. These results were consistent with a previous study by Burgess et al. (2001), who observed preferential activation of the right PHC during intentional retrieval of spatial contextual information, relative to retrieval of object and nonspatial source information (i.e., which person gave them an object).

In summary, although there is evidence that context plays a small but consistent role in recognition for verbal materials, less is known about the role of visual context in episodic object recognition. Studies focusing on pre-existing semantic relationships between object and context and lesion studies in monkeys suggest that context should play a significant role in episodic memory for uniquely presented objects. Neuroimaging studies of scene processing (Epstein and Kanwisher, 1998), object identification (Bar and Aminoff, 2003), and intentional retrieval of visual context information (Hayes et al., 2004) suggest that the medial temporal lobes, most likely the PHC, may be involved in visual context effects mediating episodic object recognition, although no study we are aware of has directly addressed this issue.

The current set of experiments was motivated by three main questions. First, does context reliably facilitate episodic object recognition in the same way that it influences episodic word recognition and semantic object identification? To maximize the possibility of observing a context effect in object recognition, we utilized unique photographs of stimuli where the object and its visual background were semantically related and maintained expected spatial relationships—the same stimulus properties shown to influence object identification. Second, after establish-
ing an effect of context on object recognition, what are the encoding and retrieval conditions that influence context-dependent object recognition? Obviously we cannot replicate the extensive set of variable manipulations that exist for syllables, words, and other verbal materials. For this first set of experiments, therefore, we chose to focus on variables that might provide insight into the degree to which objects are bound to the visual scenes in which they are presented. Third, using fMRI, we explored which regions of the medial temporal lobes mediate visual context effects in episodic object recognition.

To address these questions, Experiment 1 presented objects in visually rich and unique contexts and then presented the object on a white background at test. Experiment 2 investigated the influence of intentional, as opposed to incidental, encoding instructions. Experiment 3 manipulated the familiarity of the object targets relative to distracters, and Experiment 4 investigated whether adding a visually rich, but novel, context at test is disruptive to recognition as removing a previously presented rich context. Finally, using fMRI, Experiment 5 investigated the neural correlates of the context effects observed in the behavioral studies.

**EXPERIMENT 1**

**Methods**

**Participants**

The participants in all of the experiments were healthy University of Arizona undergraduates who gave written informed consent and received course credit as compensation. All experimental procedures were approved by the University of Arizona institutional review board. Experiment 1 included 24 participants (12 males, 12 females; mean age, 19.6 yr).

**Materials**

Pictures of unique objects (240), each in a unique context, were taken in six houses and two department stores. Images were edited in Adobe Photoshop to create two types of stimuli for each object-in-context image: an object image (the object alone on a white background) and a scene image (the object in a visually rich scene). Importantly, the object image was cut directly from the scene image, using Photoshop so that the size and perspective of the object were exactly the same in the object and scene image, resulting in four study–test conditions (see Fig. 1):

1. SCENE.SCENE: An object in a rich, naturalistic scene was presented during study. At test, participants were shown the same object in the same scene.
2. OBJECT.OBJECT: An object on a white background was presented during study. At test, participants were shown the same object on the same white background.
3. SCENE.OBJECT: An object in a visually rich scene was presented during study. At test, participants were shown the same object, although this time the object was presented on a white background.
4. CONTROL: The stimuli in this condition were Fourier transformed images of stimuli from the earlier conditions overlaid with an “X” or an “O.” A control condition was included in the behavioral experiments to increase the likelihood that behavioral results would replicate in the fMRI study (Experiment 5). In our fMRI study, this condition was designed to control for presentation of a visual stimulus, decision making, and motor response. Further discussion of the control stimuli will not be included in the behavioral experiments.

**Test materials**

The test list consisted of 220 trials: 40 target objects presented on a white background that had been presented on a white background during study (OBJECT.OBJECT), 40 target objects presented in the same visually rich context as during study (SCENE.SCENE), 40 target objects presented on a white background that were presented in a visually rich context during study (SCENE.OBJECT), 40 object lures (novel object on a white background), 40 scene lures (novel object presented in a novel scene), and 20 CONTROL trials. The object lures and scene lures were similar to the targets, i.e., indoor pictures of household objects (see Fig. 1). The 220 trial test list was subdivided into blocks of 10. Presentation order of blocks and items within blocks was randomized by DMDX. There were two test lists, counterbalanced across participants. Items from Study List B served as lures for Test List A, and items from Study List A served as lures for Test List B.

**Procedure**

After giving informed consent, participants received incidental encoding instructions. Participants were told that the study focused on the perception of objects. They were informed that they would see a set of objects, and that they would be asked to make a price judgment about each one. If the object was presented in a scene, participants were asked to make a price judgment about the object in the center of the picture. Participants responded with a left mouse button press if they thought the object cost less than $25 or with a right mouse button press if they thought the object cost more than $25. For control trials,
participants made a left mouse button press if an “X” appeared on the screen or a right mouse button press if an “O” appeared on the screen. Experimental stimuli were presented for 3 s and control stimuli were presented for 1.5 s with an intertrial interval of 1.5 s for all trials. Stimuli were presented on a PC computer with a 17-in. color monitor using DMDX (version 2.4.06; Forster and Forster, 2003). Button presses and response times were recorded.

To confirm correct button mapping and that the participants understood the instructions, a practice session was presented first, consisting of 12 control trials, and 12 price judgment trials (six objects on a white background and six objects in a scene). During the practice session, instructions were presented on the computer screen preceding each trial. Participants were informed that the instructions would not appear during the experiment proper and that they should memorize the button mapping. After the practice session, participants completed the encoding session as described earlier. Participants were not aware that a recognition test would follow.

Immediately following the encoding session, participants received a surprise yes–no recognition test. They were instructed to press the left mouse button on a trial if the object was old, even if they had previously seen the object in a naturalistic context, and to press the right mouse button if the object was new. Presentation rate and intertrial interval were identical to the encoding session. After the recognition test, participants were debriefed.

### Results and Discussion

For all behavioral data, unless indicated otherwise, a one-way repeated measures analysis of variance (ANOVA) was used to compare recognition hit rates in the three experimental conditions (OBJECT.OBJECT, SCENE.SCENE, and SCENE.OBJECT), and followed up with paired t-tests using SPSS 12.0 for Windows. False alarms rates (object lures, scene lures) were compared using a paired t-test. The alpha level for all significance tests was set at 0.05. For each experiment, recognition analyses were also conducted using d-prime. d-Prime scores were computed by adding 0.5 to the frequency of the hits and false alarms and dividing by \( N + 1 \), where \( N \) is the number of old or new trials, as described by Snodgrass and Corwin (1988). In all cases, the results of d-prime analyses were consistent with analyses of hit rates alone. d-Prime scores by experiment and by condition are presented in Fig. 2.

Results of Experiment 1 are presented in Table 1. Under incidental encoding instructions, recognition hit rates were significantly different across the three target conditions, \( F(2,46) = 24.52, P < 0.001 \). Importantly, a CSD was observed: recognition performance was significantly lower in the SCENE.
OBJECT condition relative to the OBJECT.OBJECT and SCENE.SCENE conditions, \( t(23) > 4.99, P < 0.001 \). Hit rates were equivalent in the OBJECT.OBJECT and SCENE.SCENE condition, \( t(23) < 1, \text{ns} \). The mean false alarm rate was 11.2%. There was no difference in the false alarm rate between scene lures (11.1%) and object lures (11.4%), \( t(23) < 1, \text{ns} \).

In this study, the binding of object and context information may have been facilitated by the presentation of objects in visually rich, naturalistic scenes (Bar and Ullman, 1996; Murnane et al., 1999). Because the CSD in the present experiment was observed under incidental encoding conditions, we hypothesized that the binding of object and context occurred automatically. Experiment 2 was designed to strengthen this assertion, by determining whether intentional encoding instructions would enhance recognition performance in the SCENE.OBJECT condition and eliminate the CSD effect. According to Hasher and Zacks (1979), automatic, as opposed to effortful, operations should not be influenced by intention to remember. Therefore, if the binding of object and context information is indeed an automatic operation, intentional encoding instructions to disregard the scene should not attenuate the CSD effect.

### Experiment 2

#### Methods

**Participants**

Experiment 2 included 24 participants (9 males, 15 females; mean age, 19.2 yr).

**Materials and procedure**

Study and test materials, lists, and trial conditions were identical to Experiment 1. The procedure was identical to Experiment 1, except that in Experiment 2, the study phase instructions emphasized intentional encoding and highlighted the importance of the objects. Participants were informed that they would be tested on their memory for the objects that they were rating. They were instructed to “Pay attention to the object in the center of the picture, because you will be tested only on your memory for the object.” To further accentuate the importance of the objects, participants were shown diagrams illustrating the different study–test trial types, and were informed that they would be tested on some objects that were removed from their prior context. At test, the instructions emphasized once again that they were to respond “old” to any object previously seen, regardless of whether or not it was presented in its original context. Participants completed the practice phase with the intentional encoding instructions, followed by a practice recognition test composed of four trials from each experimental condition. After completion of the practice session, the intentional encoding instructions were repeated, and the study, test, and debriefing sessions were completed following procedures described in Experiment 1.

#### Results and Discussion

The results for Experiment 2 are presented in Table 2. Recognition hit rates differed significantly across the three target conditions, \( F(2,46) = 26.44, P < 0.001 \). Despite receiving intentional encoding instructions to focus solely on the objects, a CSD was observed, as object recognition was impaired when the context changed between study and test. Consistent with Experiment 1, accuracy was significantly lower in the SCENE.OBJECT condition relative to the OBJECT.OBJECT and SCENE.SCENE condition, \( t(23) > 5.62, P < 0.001 \). Recognition was equivalent in the OBJECT.OBJECT and SCENE.SCENE conditions, \( t(23) < 1, \text{ns} \).

In this study, the binding of object and context information may have been facilitated by the presentation of objects in visually rich, naturalistic scenes (Bar and Ullman, 1996; Murnane et al., 1999). Because the CSD in the present experiment was observed under incidental encoding conditions, we hypothesized that the binding of object and context occurred automatically. Experiment 2 was designed to strengthen this assertion, by determining whether intentional encoding instructions would enhance recognition performance in the SCENE.OBJECT condition and eliminate the CSD effect. According to Hasher and Zacks (1979), automatic, as opposed to effortful, operations should not be influenced by intention to remember. Therefore, if the binding of object and context information is indeed an automatic operation, intentional encoding instructions to disregard the scene should not attenuate the CSD effect.

#### Table 1.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Hit Rate</th>
<th>False Alarm Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCENE.SCENE</td>
<td>84.1 (1.6)</td>
<td>11.3 (2.5)</td>
</tr>
<tr>
<td>OBJECT.OBJECT</td>
<td>84.0 (2.1)</td>
<td>11.3 (2.5)</td>
</tr>
<tr>
<td>SCENE.OBJECT</td>
<td>69.3 (3.0)</td>
<td>11.0 (2.5)</td>
</tr>
<tr>
<td>FA OBJECTS</td>
<td>11.3 (2.5)</td>
<td></td>
</tr>
<tr>
<td>FA SCENES</td>
<td>11.0 (2.5)</td>
<td></td>
</tr>
</tbody>
</table>
and no significant interaction between instruction and target condition, $F(1,46) < 1$, ns.

The results of Experiments 1 and 2 suggest that the binding of an object and context is relatively automatic, at least to the extent that participants are unable to consciously disregard the scene during the study episode. However, the results could also be attributed to differential levels of target familiarity across the three critical conditions. The SCENE.SCENE and OBJECT.OBJECT trials provided what Tulving (1983) referred to as “copy cues”—items that were identical at study and test and should, therefore, be highly familiar. Lures were completely novel items, and thus should have very low familiarity. On the other hand, trials in the SCENE.OBJECT condition were neither highly familiar nor completely novel, and should fall somewhere in between copy cues and lures in terms of familiarity because a portion of the stimulus was familiar (the object), whereas the remaining portion of the context was novel (white background). If participants were relying on familiarity to make their recognition judgments, particularly if they set a decision threshold for the “old” judgment rather high, then they would endorse fewer objects in the SCENE.OBJECT condition than in the other two conditions.

In Experiment 3, we attempted to reduce the effect of the ambiguous familiarity of the SCENE.OBJECT condition by presenting each set of target items in separate retrieval blocks. Thus, at retrieval, SCENE.OBJECT targets would be mixed only with novel object lures. Because these moderately familiar objects were mixed only with completely novel objects, participants may be more likely to distinguish targets from lures.

**Results and Discussion**

The results, presented in Table 3, suggest that the CSD was not attenuated by our attempt to reduce the effect of familiarity of the SCENE.OBJECT targets on recognition. There was a significant difference in hit rates across the three conditions, $F(2,46) = 19.58, P < 0.001$. The recognition hit rate was significantly lower in the SCENE.OBJECT condition than in the OBJECT.OBJECT and SCENE.SCENE conditions, $t(23) > 4.79, P < 0.001$, whereas the hit rates were similar in the OBJECT.OBJECT and SCENE.SCENE condition, $t(23) < 1$, ns.

Unlike the previous two experiments, participants false alarmed more frequently to object lures (7.8%) than to scene lures (3.8%), $t(23) = 2.86, P < 0.01$.

The significant difference in the false alarm rates between object lures and context lures suggested an analysis of corrected recognition scores. The false alarm rate for the appropriate condition was therefore subtracted from the hit rate for the targets presented in the same block. A one-way repeated measures ANOVA once again showed a significant difference in recognition across conditions, $F(2,46) = 26.09, P < 0.001$. Correcting for differential false alarm rates, all three conditions now differed significantly from one another, $t(23) > 3.7, Ps < 0.05$. In contrast to Experiments 1 and 2, targets in the SCENE.SCENE condition were better recognized (80.7%) than targets in the OBJECT.OBJECT condition (72.7%), which in turn were better recognized than targets in the SCENE.OBJECT condition (64.3%).

Taken together, the results of Experiments 1–3 suggest that the binding of an object and context is relatively automatic, impervious to the intentions of the participant, and the CSD effect cannot be attributed to differential familiarity. However, thus far we have only tested the CSD by switching from a rich context at study (object in a scene) to a relatively sparse context at test (object on a white background). As discussed earlier, the formation of a bound representation of object and context may be particularly facilitated when an object is encountered in a visually rich context, such as a scene (Murnane et al., 1999; Bar, 2003, 2004). In contrast, an object presented in a sparse context, such as a plain white background, may result in a representation that is more flexible and can be utilized later on to identify objects in any context. Alternatively, objects may be equally bound with the context, regardless of whether it is visually rich or sparse, and in both cases, changing this context from study to test will have

**TABLE 2.**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Recognition Hit (%)</th>
<th>Standard Error of Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCENE.SCENE</td>
<td>80.4 (3.2)</td>
<td></td>
</tr>
<tr>
<td>OBJECT.OBJECT</td>
<td>79.6 (2.2)</td>
<td></td>
</tr>
<tr>
<td>SCENE.OBJECT</td>
<td>67.7 (2.8)</td>
<td></td>
</tr>
<tr>
<td>FA OBJECTS</td>
<td>9.8 (1.1)</td>
<td></td>
</tr>
<tr>
<td>FA SCENES</td>
<td>8.9 (1.3)</td>
<td></td>
</tr>
</tbody>
</table>

**EXPERIMENT 3**

**Methods**

**Participants**

Experiment 3 included 24 participants (9 males, 15 females; mean age, 19.3 yr).

**Materials and procedures**

Study materials, trials, and study lists were the same as in Experiment 1. Test materials, trials, and conditions were the same as Experiment 1, but the test lists were constructed differently. In this experiment, testing occurred in three blocks, with order of presentation counterbalanced across subjects. Block A consisted of 40 OBJECT.OBJECT targets, 27 object lures, and 13 control trials. Block B consisted of 40 SCENE.SCENE targets, 27 context lures, and 13 control trials. Block C consisted of 40 SCENE.OBJECT trials, 27 object lures, and 13 control trials. The within-block target to lure ratio was consistent with the target to lure ratio in Experiments 1 and 2 (3 targets: 2 lures). Procedures and instructions to the participants were identical to Experiment 1 with the exception of the organization of test materials described earlier.
an equivalent disruptive effect on recognition. Experiment 4 was designed to test this notion. In this study, we reversed the SCENE.OBJECT condition such that objects were presented during study on a white background, and then presented in a visually rich scene at test (OBJECT.SCENE). If the representation derived at study binds the object and context automatically, and does not require a visually rich context, then the CSD effect should be observed. If, however, the presentation of objects in a visually rich context is necessary for binding of object and context information, then we would expect equivalent recognition levels in all study–test conditions, and elimination of the CSD effect.

EXPERIMENT 4

Methods

Participants

Experiment 4 included 24 participants (9 males, 15 females; mean age, 19.4 yr).

Materials and procedure

Study and test materials, trial types, and trial conditions were the same as Experiments 1 and 2, except the stimuli in the SCENE.OBJECT condition were reversed: study trials for this condition consisted of an OBJECT on a white background, and the test trials consisted of the same object in a novel and visually rich scene context. This condition will be referred to as the OBJECT.SCENE condition. The procedure in Experiment 4 was identical to Experiment 1, with the exception of the changes in materials noted earlier.

Results

As in previous experiments, the recognition hit rates across the three experimental conditions differed significantly, $F(2, 46) = 32.34, P < 0.001$ (see Table 4). Once again, a CSD effect was observed, as recognition accuracy was lower in the OBJECT.SCENE condition than in the OBJECT.OBJECT and SCENE.SCE-NE conditions, $t(23) > 6.46, P < 0.001$. Recognition hit rates were statistically equivalent in the OBJECT.OBJECT and SCENE.SCENE conditions, although there was a trend towards lower recognition hit rates for the OBJECT.OBJECT trials compared with SCENE.SCENE trials, $t(23) = 1.92, P = 0.07$. The false alarm rates for scene lures (10.4%) and object lures (9.1%) were similar, $t(23) < 1$, ns, consistent with Experiments 1 and 2 where we also employed a mixed-list paradigm.

An important question is whether the CSD effect was as large when the objects were originally encoded in the presence of a sparse context (Experiment 4) as when they were originally encoded in a visually rich context (as in Experiment 1). We therefore compared recognition scores across Experiments 1 and 4, in which the only difference in the paradigms was the order of presentation of the stimuli in the context-change condition: SCENE.OBJECT in Experiments 1–3 and OBJECT.SCENE in Experiment 4. We refer to both these conditions as “context-shift” conditions. Recognition hit rates were compared with a mixed factor ANOVA, comparing experiment (Experiment 1 vs. Experiment 4) as the between-subjects factor and trial condition (OBJECT.OBJECT, SCENE.SCENE, CONTEXT-SHIFT) as the within-subjects factor. The ANOVA revealed no effect of experiment on recognition performance, $F(1, 46) = 3.36$, ns, and no significant interaction between experiment and trial condition, $F(1, 46) < 1$, ns. Thus, the CSD effect was comparable regardless of whether the objects were studied in a sparse context (14.5%) or a visually complex context (14.8%).

Discussion

To summarize across the four behavioral experiments, a CSD effect was reliably observed in each of the four object recognition experiments. The effect was robust, with average decreases in rec-

### TABLE 3.

Mean recognition hit rate, false alarm (FA) rate, and corrected recognition scores (hits - FA) (standard error of mean) observed in Experiment 3. Note that target conditions were presented in separate test lists.

<table>
<thead>
<tr>
<th>Condition</th>
<th>SCENE.SCENE</th>
<th>OBJECT.OBJECT</th>
<th>SCENE.OBJECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Recognition hits</td>
<td>84.6 (1.5)</td>
<td>82.1 (2.3)</td>
<td>70.0 (2.4)</td>
</tr>
<tr>
<td>2. False alarms</td>
<td>3.9 (1.1)</td>
<td>9.4 (1.4)</td>
<td>5.7 (1.3)</td>
</tr>
<tr>
<td>3. Corrected rec.</td>
<td>80.7 (2.0)</td>
<td>72.7 (2.5)</td>
<td>64.3 (3.0)</td>
</tr>
</tbody>
</table>

### TABLE 4.

Mean Recognition Hit and False Alarm Rates (Standard Error of Mean) Observed in Experiment 4 Where the Scene Shift Condition (OBJECT.SCENE) Consisted of Objects Studied in a White Background and Tested in Visually Rich, Unique Contexts

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean Recognition Hit Rate</th>
<th>Mean FA Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCENE.SCENE</td>
<td>79.0 (3.0)</td>
<td>10.4 (1.8)</td>
</tr>
<tr>
<td>OBJECT.OBJECT</td>
<td>75.0 (3.1)</td>
<td>9.1 (1.4)</td>
</tr>
<tr>
<td>OBJECT.SCENE</td>
<td>64.5 (3.1)</td>
<td>10.4 (1.8)</td>
</tr>
<tr>
<td>FA OBJECTS</td>
<td>9.1 (1.4)</td>
<td></td>
</tr>
<tr>
<td>FA SCENES</td>
<td>10.4 (1.8)</td>
<td></td>
</tr>
</tbody>
</table>

_Hippocampus_ DOI 10.1002/hipo
To understand the potential binding mechanisms underlying the CSD, we should consider how items and contexts are encoded and retrieved. Remembering the occurrence of items and contexts may involve two mechanisms: one in which item and context are “fused” into a single representation and/or another mechanism, in which the representation of item and context exist separately but are “linked.” Different theoretical perspectives have referred to these two mechanisms, respectively, as integration and association (Johnson and Chalfonte, 1994; Murnane et al., 1999), elemental and configural association (Rudy and Sutherland, 1995), or blended and relational (Cohen and Eichenbaum, 1993; Moses and Ryan, 2006), although the conceptual distinctions between the two processes appear to be relatively similar in all cases. For our current purposes, we will refer to these mechanisms as “integration” and “association.”

Although it is tempting to say that the large CSD effect observed here suggests a strongly integrated representation, in fact, both integrative and associative mechanisms predict that removal of the original context at test will negatively impact object recognition. If the object and context are integrated, then the object presented without the context will be treated as a different item, making it more difficult to recognize when it is removed from its original learning context. If the object and context are associated, the context can be used as a cue for retrieving the associated object representation, making recognition more efficient and more accurate. When the cue is removed, recognition should become more difficult. The current behavioral experiments cannot distinguish between integrative or associative processing as the mechanism underlying the CSD. This issue will be addressed further in Experiment 5, an fMRI study.
Related to the current set of behavioral studies, if the CSD is mediated by association (relational memory), then greater activation in the hippocampus proper should be observed during successful recognition in the SCENE.OBJECT condition when compared with the OBJECT.OBJECT condition. Such a finding would suggest that object and context were stored separately, and that successful recognition depends upon hippocampal activation, reflecting the flexible use of the object representation. On the other hand, if the CSD is mediated by a blended representation that cannot be utilized in a flexible manner, then successful recognition may be mediated by non-hippocampal regions.

Given that the PHC has been identified with scene processing (Epstein and Kanwisher, 1998), retrieval of visual semantic association (Bar and Aminoff, 2003; Epstein and Higgins, 2006), and intentional episodic retrieval of visual contextual information (Burgess et al., 2003; Hayes et al., 2004), PHC would be a likely candidate to mediate the CSD. To investigate these issues, we replicated Experiment 1, while collecting fMRI data using a rapid event-related design. Importantly, participants were scanned during both encoding and recognition. This design allowed us to compare activity associated with memory hits and misses (for review, see Paller and Wagner, 2002), an advantage allowed us to compare activity associated with memory hits and misses (for review, see Paller and Wagner, 2002), an advantage allowed us to compare activity associated with memory hits and misses (for review, see Paller and Wagner, 2002), an advantage allowed us to compare activity associated with memory hits and misses (for review, see Paller and Wagner, 2002), an advantage allowed us to compare activity associated with memory hits and misses (for review, see Paller and Wagner, 2002).

Methods

Participants

Experiment 5 included 23 participants, screened for contraindications to MRI. Participants were oriented to the fMRI scanning procedure prior to the imaging session. Three participants were excluded from neuroimaging analyses (two on the basis of their recognition memory performance, which did not differ from chance, and one due to excessive head motion). The remaining 20 participants were included in all analyses (6 males, 14 females; mean age = 20.2 yr, mean education = 12.6 yr).

Materials and procedure

Materials were identical to Experiment 1, with the exception that additional CONTROL trials were included for appropriate modeling of the hemodynamic response. The procedure was also identical to Experiment 1, with the following protocol changes made to accommodate collection of fMRI data. The practice encoding session took place outside the scanner. Participants were then placed supine on the MRI table, fitted with high-resolution goggles, earplugs, and earphones (Resonance Technologies, Los Angeles, CA), and had their heads stabilized with cushions. The participants were moved into the bore of the scanner, and sagittal localizer and high resolution structural scans were collected (see Image Acquisition section). Next, participants completed the study session (incidental encoding), while undergoing functional scanning. Stimulus order and presentation duration were determined by optseq2, a software program designed to maximize statistical efficiency in rapid event-related fMRI analyses (http://surfer.nmr.mgh.harvard.edu/optseq/; Dale, 1999; Dale et al., 1999). Experimental stimuli were presented for 3 s and control stimuli were jittered such that presentation duration ranged from 2.25 to 6.75 s with an intertrial interval of 1.5 s for all encoding and retrieval trials. Therefore, the duration of experimental stimuli were held constant, but the onset of experimental trials were jittered to facilitate deconvolution of the hemodynamic response. Stimuli were presented via high resolution goggles using a PC with DMDX (version 2.4.06; http://www.u.arizona.edu/~jforster/dmdx.htm; Forster and Forster, 2003), a stimulus-presentation program. Button presses and response times were recorded with a mouse (modified for use in the scanner) held in the participant’s hand. After the episodic recognition test, T1 structural scans were collected, and participants were removed from the scanner and debriefed.

Image acquisition

Images were collected on a General Electric 3.0 Tesla HD Signa Excite long bore scanner (Milwaukee, WI), equipped with Optimized ACGD Gradients (40, 150 mT/m/ms slew rate running 12× software). Total scan time was ~1 hr. A sagittal localizer was collected to align a high resolution SPGR series (1.5-mm sections covering whole brain, interscan spacing = 0, matrix = 256 × 256, flip angle = 30, TR = 22 ms, TE = 3 ms, FOV = 24 cm) that was used to overlay functional images for coregistration in MNI space and anatomical localization. Following acquisition of the high-resolution anatomical images, we acquired whole-brain functional images parallel to the anterior commissure–posterior commissure plane using a single-shot spiral in/spiral out sequence (Glover and Law, 2001; direction = superior to inferior, matrix = 64 × 64, FOV = 24 cm, TR = 2.040 ms, TE = 30 ms, sections = 31, thickness = 3.8 mm, interscan spacing = 1). Functional scanning lasted ~30 min, and occurred in 2 runs (study = 334 repetitions plus 10 repetitions that were discarded to allow MR signal to normalize, test = 516 repetitions plus five repetitions that were discarded to allow MR signal to normalize). After completion of functional scanning, T1-weighted images (direction = superior to inferior, matrix = 256 × 256, FOV = 24, TR = 500 ms, TE = 17 ms, sections = 31, thickness = 3.8 mm, interscan spacing = 0) were collected in the exact orientation as the functional images.

Image processing and analysis

Using SPM2 (http://www.fil.ion.ucl.ac.uk/spm/ Wellcome Department of Imaging Neuroscience; Frackowiak et al., 1997), images were corrected for asynchronous slice acquisition (slice timing: reference slice = 15, TA = 1.97), realigned using a 4th degree b-spline interpolation, coregistered using a trilinear reslice interpolation method, normalized to MNI space, and smoothed (7-mm Gaussian kernel). The hemodynamic response for each trial was modeled using a canonical hemodynamic response function. Serial correlations were estimated using an autoregressive AR (1) model. Data were high pass filtered using a cutoff of 128 s, and global effects were removed (session specific grand mean scaling, global scaling). To identify regions of significant activation, group random-effects analyses were performed.

Hippocampus DOI 10.1002/hipo
alpha level for all contrasts was set at an uncorrected $P$-value of .005 with a minimum cluster size of five contiguous voxels.

For follow-up analyses, conjunction analyses were performed using an inclusive mask, which functions as a logical AND operator. For example, we assessed the overlap of regions involved in perceptual processing of scenes (i.e., Encoding: Scenes > Objects) AND retrieval success for the context shift condition (i.e., SCENE.OBJECT hits > SCENE.OBJECT misses). For a voxel to be considered active, it must surpass statistical thresholds for both individual contrasts. Because the activation associated with perceptual processing of scenes is robust, we set a more stringent criteria for this contrast ($P < 0.005$) with a minimum cluster size of five contiguous voxels. Comparing the effects of retrieval success (e.g., hits vs. misses) required the detection of more subtle effects, and therefore a somewhat less stringent $P$-value was used ($P < 0.01$) with a minimum cluster size of five contiguous voxels.

Figures were created using MRicro (version 1.39; http://www.mricro.com; Rorden and Brett, 2000). Brain activations were overlaid on the International Consortium of Brain Mapping single subject MRI template (http://www.loni.ucla.edu/ICBM/Downloads/Downloads_ICBMtemplate.shtml). Surface cortical activations were rendered using a depth value of 10 voxels. Thus, for an activation to be projected onto the cortical surface, it had to be located within 10 voxels (20 mm) of the cortical surface.

Results

Behavioral results

Hit and false alarm rates are presented in Table 5. Under incidental encoding instructions, recognition hit rates were significantly different across the three target conditions, $F(2,38) = 35.34, P < 0.001$. Importantly, consistent with the previous behavioral experiments, a CSD was again observed. Recognition performance was significantly lower in the SCENE.OBJECT condition relative to the OBJECT.OBJECT and SCENE.SCENE conditions, $t$ (19) > 6.19, $P < 0.001$. Hit rates were equivalent in the OBJECT.OBJECT and SCENE.SCENE condition, $t$ (19) < 1, ns. The mean false alarm rate was 12.2%. There was no difference in the false alarm rate between scene lures (12.4%) and object lures (12.0%), $t$ (19) < 1, ns.

Because response times can impact interpretation of the neuroimaging data, response times for each trial type were compared (see Fig. 4). One participant was excluded from the response time analyses because the participant did not false alarm to any object lures. A repeated measures ANOVA comparing memory judgment (correct, incorrect) and condition (OBJECT.OBJECT, SCENE.SCENE, SCENE.OBJECT, Object Lures, Scene Lures) revealed a main effect for memory judgment, $F(1,18) = 18.17, P < 0.001$, with faster response times for correct ($M = 1.122$ ms, SEM = 44) than incorrect ($M = 1.420$ ms, SEM = 95) responses. There was also a main effect of condition, $F(4,72) = 3.04, P < 0.05$, and pairwise comparisons revealed that participants took longer to respond to scene lures than any other condition, $P's < 0.05$. The memory judgment by condition interaction was also significant, $F(4,72) = 3.43, P < 0.05$. Pairwise comparisons revealed no differences in response time for incorrect trials, ns. For target memory conditions, pairwise comparisons indicated that response times were fastest for OBJECT.OBJECT correct trials, followed by SCENE.SCENE correct trials, and then SCENE.OBJECT correct trials, $P's < 0.05$. Thus, participants were slowest to respond to the SCENE.OBJECT correct trials, which in fact did not differ in response times from correct rejections (Object Lures correct and Scene Lures correct trials), ns.

fMRI results

Scene processing. We contrasted encoding of scenes versus objects, and observed robust activation in posterior brain regions typically associated with processing complex visual stimuli, including bilateral PHC, fusiform gyrus, and lateral parietal cortex (see Fig. 5A). The results of this random effects analysis were used as a mask for subsequent conjunction analyses.
A conjunction analysis was used to investigate encoding success activity for the SCENE.OBJECT condition, comparing activation at encoding for those SCENE.OBJECT trials that were later recognized versus those that were forgotten (referred to as a subsequent memory analysis). The results yielded a subsequent memory effect in right PHC. Activation was also observed in a small region of left PHC, although the majority of the cluster of activation resided within right PHC.

**FIGURE 5.**

A: Activation observed when contrasting encoding of scenes (object in scene) versus objects [image coordinates: MNI: 30 -48 -8; Tal: 29 -47 4]. Massive bilateral activation was observed in posterior brain regions, including the following regions: PHC, fusiform and lingual gyri, inferior, middle and superior occipital gyri, cuneus, precuneus, and inferior portions of the superior parietal lobe. B: Subsequent memory effects in PHC, right [MNI: 28 -42 -6; Tal: 28 -41 -3] greater than left [MNI: -32 -51 -3; Tal: -32 -50 0]. Increased PHC activation during encoding of SCENE.OBJECT trials was associated with subsequent memory for the object when it was later presented on a white background. C: Retrieval success activity in SCENE.OBJECT condition. Preferential activation of right PHC [MNI: 30 -41 -9; Tal: 30 -41 -5] was observed in SCENE.OBJECT hits > SCENE.OBJECT misses. D: Bilateral PHC activation during recognition hits of SCENE.OBJECT relative to OBJECT.OBJECT [right PHC, MNI: 28 -36 -14; Tal: 28 -35 -10; left PHC, MNI: -28 -40 -10; Tal: -28 -40 -7]. The only difference between the conditions was that SCENE.OBJECT trials had been encoded in a naturalistic scene, whereas OBJECT.OBJECT trials had been encoded on a white background.

**Encoding success activity.** A conjunction analysis was used to investigate encoding success activity for the SCENE.OBJECT condition, comparing activation at encoding for those SCENE.OBJECT trials that were later recognized versus those that were forgotten (referred to as a subsequent memory analysis). The results yielded a subsequent memory effect in right PHC. Activation was also observed in a small region of left PHC, although the majority of the cluster of activation resided within right PHC.
left fusiform gyrus (see Fig. 5B). As seen in Figure 6, within right PHC, activation was greater at encoding for subsequent hits than misses in the SCENE.OBJECT condition, whereas there was no difference between subsequent hits and misses in this region for OBJECT.OBJECT trials. A random effects analysis directly investigating encoding success activation (hits vs. misses) in the OBJECT.OBJECT condition resulted in activation in bilateral fusiform gyrus, inferior temporal gyrus, and the lateral occipital complex, although activation in PHC was not observed. Thus, activation in right PHC was associated with successful encoding of items in the SCENE.OBJECT and SCENE.SCENE conditions, but it was not associated with successful encoding of OBJECT.OBJECT trials. Other regions that were significantly active for subsequent memory of SCENE.OBJECT trials included bilateral fusiform gyrus, bilateral middle occipital gyrus, bilateral retrosplenial cortex, and right precuneus.

Retrieval success activity. The second main comparison investigating the neural correlates of the CSD contrasted retrieval success (hits vs. misses) in the SCENE.OBJECT condition, using the scene processing contrast as an inclusive mask for the conjunction analysis. Thus, in this analysis, both conditions consisted of trials in which a studied object had undergone a context shift—from a naturalistic scene at encoding to a white background at retrieval. The only difference between the conditions was whether the participant correctly recognized the object as an old item or not. Results indicated activation primarily in right PHC extending into the fusiform gyri (see Fig. 5D). Additional activation was observed in the parieto-occipital sulcus, and bilateral cuneus. It is important to note once again that, in this comparison of SCENE.OBJECT hits > OBJECT.OBJECT hits, objects were presented on a white background in both conditions at recognition. In both conditions, participants correctly identified the object as “old.” Thus, the only difference between the two conditions was prior visual experience with the object—that is, whether it was originally encoded in a naturalistic scene or on a white background. The results of this analysis suggest that PHC is not only involved in processing of visual scene information, but that PHC may play an important role in the reinstatement of visual contextual information associated with an object, and that reinstatement will facilitate episodic recognition memory, even when the encoded information is not perceptually reinstated at retrieval.

Discussion

To summarize, the right PHC appears to play a critical role in the CSD effect. Encoding success activity was observed predominantly in the right PHC in the SCENE.OBJECT condition, whereas encoding success activity was not observed in PHC during the OBJECT.OBJECT condition. Right PHC was also associated with retrieval success activity in the SCENE.OBJECT condition, whereas retrieval success activity was not observed in this region in OBJECT.OBJECT condition. Bilateral PHC activation was observed during successful retrieval of SCENE.OBJECT trials relative to OBJECT.OBJECT trials, suggesting that visual contextual information may be reinstated at test, and that reinstatement may facilitate episodic object recognition. The results presented here expand the function of PHC beyond perceptual scene processing, highlighting its importance for successful episodic memory encoding and retrieval. The localization of the CSD effect to PHC suggests that integration, as opposed to association, may be the mechanism underlying the CSD. Each of these topics will be addressed in more detail later, in addition to a brief discussion of the current results in relation to Multiple Trace Theory (MTT) (Nadel and Moscovitch, 1997, 1998; Nadel et al., 2003).

Neural correlates of the CSD

Encoding success activity. In the present study, PHC was preferentially activated, right greater than left, in SCENE.OBJECT trials that were subsequently remembered relative to those that were subsequently forgotten. Previous blocked design studies have found similar activation (regardless of subsequent memory) during intentional encoding of pairs of unrelated visual stimuli in the right parahippocampal gyrus and right hippocampus (Henke et al., 1997) and bilateral parahippocampal gyrus and hippocampus (Rombouts et al., 1997; Pihlajamaki et al., 2003). A more recent event-related fMRI study observed encoding success activity in right parahippocampal gyrus associated with correct perceptual source judgments relative to semantic source
judgments, again under intentional encoding instructions (Prince et al., 2005). Thus, in each of these studies, scanning occurred under intentional encoding instructions (see also Kohler et al., 2002). The current paradigm is unique in that participants were scanned under incidental, object-focused encoding instructions (making a judgment about the price of an object). Despite the incidental nature of the encoding task and object-focused instructions, subsequent memory effects were observed predominantly in right PHC. Similar encoding success activity was not observed in PHC on the right or left for the OBJECT.OBJECT condition, highlighting the specificity of this effect for objects embedded in complex scenes, rather than singular objects. The data from the current study are consistent with a previous report of encoding success activity in bilateral PHC observed during incidental encoding of scenes (Brewer et al., 1998), although in this case participants scene recognition was later assessed, as opposed to recognition for objects previously presented in a scene.

Although recognition memory was not assessed, a related fMRI study used an adaptation paradigm during passive viewing of naturalistic pictures (objects in scenes; Goh et al., 2004). Briefly, quartets of picture were rapidly presented in succession, with a component of the picture (object, scene, or both) repeated. Reduced activation, i.e., adaptation, was observed in bilateral fusiform gyrus when the object was repeated, bilateral parahippocampal gyrus when the scene was repeated, and bilateral parahippocampal gyrus and right anterior hippocampus when the object and scene were repeated. The pattern of results suggests that posterior portions of the parahippocampal gyrus may be involved with the automatic binding of object and scene information. Furthermore, consistent with the present study, Goh et al. (2004) found object-related effects in fusiform gyrus, as opposed to PHC. Data from the current study extend the results of Goh et al. (2004), as the present results indicate that differences in activation within these regions during incidental encoding are associated with differential subsequent episodic memory performance. Both the behavioral data and the fMRI encoding data from the present study, and the results of Goh et al. (2004), are consistent with the notion of obligatory and automatic binding of consciously apprehended information, in this case object and scene (Moscovitch, 1994).

Retrieval success activity. In the SCENE.OBJECT condition, at retrieval, greater activation was observed in right PHC for hits relative to misses. In a previous experiment in our laboratory using similar materials, we observed preferential activation of PHC during intentional retrieval of spatial contextual information relative to nonspatial source information (temporal order; Hayes et al. 2004). Consistent with these results, Burgess et al. (2001) also found greater right PHC activation during spatial judgments relative to nonspatial source judgments (which character presented the participant with an object). Again, the current paradigm is unique because participants were not explicitly required to make memory judgments about the visual context—yet PHC activation was observed during successful performance on the object recognition task. The results suggest that PHC may automatically reinstate object and context information in response to a partial cue, in this case, the object alone.

Tsivilis et al. (2003) recently reported an event-related fMRI study of object and context repetition effects, but found no difference in object recognition performance when context was shifted between study and test. While Tsivilis et al. (2003) presented objects in naturalistic contexts, and then manipulated familiarity of the object and its context, the objects were manually segregated from the scene, bounded with a yellow outline, superimposed on various backgrounds, and then studied twice by the participants. Furthermore, in the identical condition, the position of the object shifted between study and test, a manipulation that has been shown to result in decreased memory performance (Hollingworth, 2006). In contrast, the current study presented visually integrated photographs of objects in related and naturalistic contexts, and strong CSD effects were observed when the context was shifted. The difference in outcomes across the two studies raises the question of what properties must be present in the scene to obtain strongly bound representations, such as the semantic relatedness of the background to the object, or the visual integration of the object with the background.

Automatic reactivation of scene information. To determine if reactivation of scene processing areas mediated successful recognition performance, we compared SCENE.OBJECT hits with OBJECT.OBJECT hits during recognition. In both conditions, objects were presented on a white background, and participants made correct memory judgments for objects in both conditions. It should be emphasized that participants were not asked to make judgments about scene information. The only difference between the conditions was the previous encoding experience, that is, whether objects were encoded in a naturalistic scene or on a white background. Robust bilateral PHC and fusiform activation was observed, the same regions involved in encoding of scene information, but only for objects that were originally encoding in a scene. The current fMRI results suggest that successful memory performance is supported by the PHC, and that it may be mediated by the automatic reinstatement (incidental retrieval) of scene information that was presented during the original encoding experience.

Previous experiments have investigated “reactivation” effects, and found overlap of regions involved with encoding and retrieval (Nyberg et al., 2000; Wheeler et al., 2000; Vaidya et al., 2002; Prince et al., 2005). The design in each of the cited studies included either intentional retrieval of source information and/or a blocked design. Use of an intentional retrieval design is problematic because selective attention to a given perceptual modality has been shown to result in activation of the same regions involved in the actual perceptual task (Cabeza and Nyberg, 2000). For example, an intentional retrieval task, such as “was a sound heard?” requires selective attention to auditory stimuli, and would therefore contaminate the reactivation results. To demonstrate that reinstatement of neural activity occurs automatically and is not confounded by selective attention or imagery, the memory test should occur under incidental retrieval conditions, where the instructions to the participants do not
emphasize the component of interest. Although we cannot con- 
firm that participants in the present study did not make an effort 
to intentionally retrieve scene information during object recogni-
tion, the use of a mixed test list and rapid event-related design, 
along with the instructions emphasizing recognition of the object 
alone, minimizes the possibility that participants adopted an 
intentional source retrieval strategy in the SCENE.OBJECT 
condition. Nevertheless, it would be interesting to know the 
degree to which participants were aware of the absence of context 
for some objects, but not others, that is, to test participants' 
knowledge of source, and whether PHC plays a similar role in 
source judgments as it does in object recognition.

Association versus integration. The absence of activation in the 
hippocampus proper in light of activation in PHC could lend 
insight into the mechanism underlying the CSD. For instance, 
O’Keefe and Nadel (1978) (see also Moses and Ryan, 2006) sug-
gest that relational memory (association in our terms) is depend-
ton hippocampal function, whereas elemental or blended 
(integrated) representations can be mediated by non-hippocam-
pal structures. According to Moses and Ryan (2006), blended 
representations do not allow for separation of individual ele-
ments included in the representation. The subsequent presenta-
tion of only one a few elements will lead to activation of the 
complete representation, or will lead to no activation at all. They 
emphasize that the elements of the representation cannot be sep-
parated. The current activation in PHC, as opposed to hippocam-
pus, therefore, suggest that object and scene information may be 
integrated, and based on the current behavioral and fMRI data, 
that this process occurs relatively automatically. Indeed, although 
Moses and Ryan refer to “activation” in reference to the cogni-
tive construct of an abstract representation, the description is ap-
licable to the current fMRI data. That is, those regions involved 
with processing of object and scene information at encoding 
were the same regions that were activated at retrieval, even when 
the objects were presented alone on a white background, suggest-
ing that they indeed cannot be separated. The results of the pres-
ent experiment are consistent with the view that objects and their 
contexts are integrated automatically into a blended representa-
tion, and that changing context from one presentation to 
another will significantly disrupt either implicit priming or 
explicit recognition. Indeed, in the current study, successful per-
formance was associated with increased encoding success activity, 
retrieval success activity, and reactivation within the PHC, par-
cipantly right PHC. Given its role in memory, one question that 
arises is why the hippocampus was not associated with encoding 
or retrieval success, or reactivation effects? First, many studies 
have implicated the PHC in scene processing (Aguirre et al., 
1996; Epstein, 2005; Epstein and Kanwisher, 1998; Epstein 
et al., 1999) and memory for visual contextual information 
(scenes, spatial-location, or visual associates; Maguire et al., 
1996; Stern et al., 1996; Brewer et al., 1998; Maguire et al., 
1998; Johnsrude et al., 1999; Mellet et al., 2000; Burgess 
et al., 2002; Duzel et al., 2003). Second, there are several patient 
udies that emphasize the importance of the right PHC in rec-
ognizing spatial-location information. Right temporal lobectomy 
patients, for example, are impaired in memory for the location 
of objects but not for the objects themselves (Pigott and Milner, 
leagues attributed this impairment in object location to damage 
to the hippocampus, although it was noted that damage 
extended beyond the hippocampus proper to include the para-
hippocampal gyrus, and indeed, the severity of object location 
impairment correlated with the extent of damage. It is possible, 
therefore, that the critical region underlying the observed deficit 
was the PHC. Bohbot et al. (1998) tested patients who had 
undergone thermo-coagulation with a single electrode along the 
amygdalo–hippocampal axis—a procedure that can limit damage 
to single structures of the medial temporal lobe—to alleviate epi-
lepsy. Patients with lesions restricted either to the right hippo-
campus proper or to the right PHC were equally impaired on an 
object location memory task. The PHC provides one of the 
major neocortical inputs to the hippocampus proper; a lesion to 
the PHC could be considered a functional lesion to the 
hippocampus.

Alternatively, the PHC and the hippocampus proper may play 
complementary, but different, roles in memory for spatial rel-
ations. Burgess et al. (2002) suggest that PHC is required for the 
iconic representation of scenes (see also Epstein et al., 2003), 
whereas the hippocampus is required only when the memory 
task requires the use of 3D space, such as when an individual is 
asked to identify a scene from an alternate point of view. The 
present results are not inconsistent with this suggestion. In the 
current design, successful recognition performance could be 
mediated by an iconic representation and would not require a 
3D representation of the object. 

The enigma posed by our data is why participants cannot rely 
on components of the representation to control behavior. It is 
possible that integrated object-context representations have pri-
iority in the queue for control of behavior, even when other com-
ponent representations (relational, for instance) also exist (cf., 
White and McDonald, 2002).

Multiple trace theory

Finally, MTT has suggested a differentiated role for MTL 
regions in human memory (Nadel and Moscovitch, 1997, 1998; 
date, MTT has focused on distinguishing the role of MTL struc-
tures in retrieval of recent versus remote episodic memory and, 
in addition, has offered some tentative statements about the role 
of MTL in episodic versus semantic memory. A central tenet of 
MTT is that the hippocampal complex is involved with the re-
trieval of contextual information. Contrary to the idea that the 
hippocampus merely acts as a “pointer” to disparate pieces of in-
formation in the cortex (Alvarez and Squire, 1994), MTT sug-
gests that the hippocampal complex not only binds disparate 
 pieces of information in neocortex, but also participates in the 
storage of contextual information, namely spatial. These context-
focused tenets of MTT were based on a comprehensive review of 
the literature suggesting that the hippocampal complex is critical 
for memory for spatial information, or cognitive maps (O’Keefe

Hippocampus DOI 10.1002/hipo
and Nadel, 1978) and the definition of episodic memory, in which the retrieval of spatial and temporal information mediates mental time travel (ability to reconstruct past events from memory; Tulving, 1973, 1983, 1985; cf., Hassabis et al., 2007). In a general sense, O’Keefe and Nadel’s conception of taxon and local memory maps roughly to Tulving’s conception of semantic and episodic memory.

The current paradigm provides preliminary evidence for expanding the theoretical scope of MTT. First, MTT posits that retrieval of contextual information is dependent upon the hippocampal complex. Additional support for this hypothesis was provided by observing encoding success activity (ESA), retrieval success activity (RSA), and reactivation within PHC during SCENE.OBJECT trials relative to trials, in which distinct contextual information was not available (OBJECT.OBJECT). Second, although MTT states that the hippocampal complex rapidly and obligatorily encodes consciously attended information, similar to Moscovitch’s description of the hippocampal system as a “stupid” processor, MTT does not currently make specific predictions about the role of regions within MTL, such as hippocampus proper and PHC, during the encoding of different types of information, e.g., object and context information. Here we provide data suggesting that a subregion of the MTL, the PHC, is involved in both incidental encoding and retrieval of scene context information.

Acknowledgments

These experiments were completed in partial fulfillment of Scott M. Hayes’ dissertation at the University of Arizona. We thank Dr. Elizabeth Glisky for her useful comments and discussion, Dr. Dianne Patterson for technical support, and Carrie McAllister and Katie Ketcham for their assistance with data collection. We would also like to thank Dr. Gary Glover for development of the spiral in/out sequences used in the current fMRI experiment and Dr. Theodore Trouard for MRI technical support.

REFERENCES


