Visual Memory and Visual Perception Recruit Common Neural Substrates

Scott D. Slotnick
Boston College

This human neuroimaging review aims to determine the degree to which visual memory evokes activity in neural regions that have been associated with visual perception. A visual perception framework is proposed to identify cortical regions associated with modality-specific processing (i.e., visual, auditory, motor, or olfactory), visual domain-specific processing (i.e., "what" versus "where," or face versus visual context), and visual feature-specific processing (i.e., color, motion, or spatial location). Independent assessments of visual item memory studies and visual working memory studies revealed activity in the appropriate cortical regions associated with each of the three levels of visual perception processing. These results provide compelling evidence that visual memory and visual perception are associated with common neural substrates. Furthermore, as with visual perception, they support the view that visual memory is a constructive process, in which features or components from disparate cortical regions bind together to form a coherent whole.

Key Words: domain specific, feature specific, fMRI, human, item memory, modality specific, PET, recall, recognition memory, working memory

Memory has been construed as a constructive process, in which features or components from disparate regions of the brain are unified into a coherent whole (Schacter, Norman, & Koutstaal, 1998; Squire, 1992). This view is reminiscent of the divergent cortical processing that occurs during visual perception (e.g., color is processed in one region whereas motion is processed in another; this is discussed in more detail below). In monkeys, memory for visual items has been found to evoke activity within the same cortical processing regions that have been associated with visual perception (Miyashita, 1993; Ungerleider, 1995).

The present review assesses the degree to which, in humans, memory for visual items evokes activity in cortical processing regions associated with visual perception, a possibility that stems from a long history of behavioral research relating memory encoding and retrieval processes (see Roediger, Gallo, & Geraci, 2002; Tulving & Thomson, 1973). Two forms of explicit memory tasks were considered in detail: item memory (a form of long-term memory), which in the laboratory refers to memory retrieval of a previously encoded item following a delay lasting minutes or longer, and working memory (short-term memory), which refers to the continued maintenance of an encoded item during a delay period (see Figure 1). This selective review focuses on functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) results, given that these methods have spatial resolutions on the order of millimeters allowing for precise spatial localization of activated neural regions (versus the centimeter spatial resolution of event-related potentials [ERPs]; see Slotnick, 2004).

The aim of this review is twofold. First, a visual perception framework is described that consists of modality-specific processing, domain-specific processing, and feature-specific processing. Second, visual item memory and visual working memory studies are considered with reference to this visual perception framework. To anticipate the results, there is a preponderance of evidence that visual-memory-related activity mirrors that of visual perception.
Figure 1: Typical Behavioral Protocols Employed in the Laboratory to Study Item Memory (Left) and Working Memory (Right).

NOTE: Abstract shapes were adapted from Slotnick and Schacter (2004). Item memory experiments begin with an encoding phase (which usually takes ten or more minutes) in which a series of items are sequentially presented and participants are instructed to remember each item. In addition to the delay imposed by the length of the encoding phase, there is usually an additional delay on the order of minutes (although this delay can last hours or days). During the item memory retrieval phase, encoded (old) shapes and new shapes are sequentially presented, and participants respond as to whether each item is "old" or "new" (which would correspond to correct responses for the two shapes shown). The term visual item memory refers to the cognitive process of memory retrieval, when visual items from the encoding phase are remembered. Working memory experiments consist of sequential trials, each with an encoding phase, a delay period, and a retrieval phase. The encoding phase can consist of a single or multiple to-be-remembered items. The delay period lasts on the order of seconds, whereas the retrieval phase consists of an old item from the preceding encoding phase or a new item, and participants respond as to whether the item is "old" or "new" (which are again the correct responses). The term visual working memory refers to the cognitive process of actively holding the visual item "in mind" during the delay period.
VISUAL PERCEPTION

Modality Specific

Although the focus here will be on visual perception, distinct cortical regions have also been associated with auditory-word perception, motor action, and olfactory perception (see Figure 2A). Auditory perception of sounds or words has been associated with bilateral activity in the superior temporal gyri extending into the inferior aspect of the Sylvian fissure (Hall et al., 2002; Price, 2000), with the addition of bilateral activity in the inferior frontal cortex during word production (Price, 2000; Stippich et al., 2003). Motor action has been associated with activity (usually contralateral to the side of movement) in the primary motor cortex within the anterior aspect of the central sulcus extending into the premotor cortex within the precentral gyrus (Picard & Strick, 2001; White et al., 1997). Olfactory perception has been associated with activity in the piriform cortex within the temporal lobe and the orbitofrontal cortex (Kareken et al., 2004; Sobel et al., 1998). Visual perception has been associated with activity within the occipital cortex extending into the inferior temporal cortex and parietal cortex (in addition to the prefrontal cortex), as is discussed in detail below.

Domain Specific

Visual spatial processing can subserve item identification or spatial localization, which can be considered distinct processing domains. Given that there has been a long history of work aimed at delineating the neural substrates of visual spatial processing in monkeys (nonhuman primates), the visual spatial cortical organization in monkeys will be briefly reviewed followed by a consideration of the results in humans. There are two visual cortical processing streams that are associated with item identification (i.e., the “what” or ventral stream) or item spatial localization (i.e., the “where” or dorsal stream); the ventral stream begins in the occipital lobe and extends from striate cortex (V1) to extrastriate cortex and into inferior temporal cortex, while the dorsal stream also begins in striate cortex and extends to extrastriate cortex and parietal cortex (Ungerleider & Mishkin, 1982). Within each visual cortical processing stream, there is a hierarchical processing organization beginning with striate cortex, which outputs to extrastriate regions, which in turn output to temporal cortex (ventral stream) or parietal cortex (dorsal stream), both of which ultimately converge on the hippocampus; there is also a massive degree of feedback within each processing stream in addition to some measure of cross-talk between processing streams (Felleman & Van Essen, 1991). There are two additions to the Felleman and Van Essen (1991) processing architecture that are of particular relevance: (a) There is evidence that the ventral processing stream extends from the inferior temporal cortex to the ventral prefrontal cortex and the dorsal processing stream extends from the parietal cortex to the dorsal prefrontal cortex (Pandya & Yeterian, 1996), and (b) the dorsal stream output converges on the parahippocampal cortex (a region that will be discussed below) before entering the hippocampus (Mishkin, Suzuki, Gadian, & Vargha-Khadem, 1997).

Human PET studies have provided evidence that is consistent with the visual spatial cortical processing architecture described in monkeys. In particular, item identification has been associated with ventral stream activity in the occipital cortex and inferior temporal cortex (in lingual and fusiform gyri) in addition to the inferior prefrontal cortex, whereas spatial localization has been associated with dorsal stream activity in the occipital cortex, superior parietal cortex, intraparietal sulcus, and inferior parietal cortex (Haxby et al., 1991, 1994; Köhler, Kapur, Moscovitch, Winocur, & Houle, 1995). As such, it can reasonably be assumed that the detailed work on monkeys indicating that the ventral stream extends from occipital cortex to inferior temporal cortex to ventral prefrontal cortex whereas the dorsal stream extends from occipital cortex to parietal cortex to dorsal prefrontal cortex is also descriptive of visual spatial processing in humans (see Figure 2B). It should be noted, however, that the prefrontal cortex distinction is considered only when reviewing visual working memory studies, where it becomes particularly relevant.

Domain specificity not only refers to the ventral and dorsal processing distinctions but can also refer to more detailed processing corresponding to different visual categories. Within the ventral temporal cortex, face processing has been associated with activity in the fusiform gyrus, whereas house processing has been associated with activity in the collateral sulcus extending into the medial fusiform gyrus and parahippocampal gyrus (see Figure 2C; Chao, Martin, & Haxby, 1999; Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999; Kanwisher, McDermott, & Chun, 1997), and these processing regions extend posteriorly into ventral occipital cortex (Ishai, Ungerleider, Martin, & Haxby, 2000). It is important to note that in all these studies, category-specific regions were responsive to nonpreferred categories to some degree. In fact, using numerous item categories (faces, houses, cats, shampoo bottles, scissors, shoes, chairs, and nonsense images), Haxby et al. (2001) showed that categories were associated with distributed and overlapping patterns of activity within the ventral temporal cortex (see also Joseph, 2001), a finding that is reminiscent of the known distributed processing architecture in monkey ventral temporal cortex (Tanaka, 1993). These results indicate that neural regions prefer-
Figure 2: Lateral and Inferior View of a Representative Participant’s Left Hemisphere Gray-White Matter Segmentation.
NOTE: To the left, the lateral view of a representative participant’s left hemisphere gray-white matter segmentation is shown. To the right, the inferior view of the same hemisphere is shown. Gyri are shown in light gray, and sulci are shown in dark gray, with those of interest labeled in white.

(A) Modality-specific processing regions, demarcated in black, include visual, auditory, motor, and olfactory. It should be noted that the ventral visual pathway traverses the inferior occipital and temporal cortex (see text) but is shown in the lateral view for illustrative purposes. Unless otherwise noted, these regions and those mentioned below occur bilaterally. (B) Domain-specific processing includes the “what” (ventral) and “where” (dorsal) pathways in addition to face and visual context processing regions. (C) Feature-specific regions process motion, color, and location. Location refers to the precise location in the visual field that evokes contralateral retinotopic activity in striate and extrastriate cortex. Stimuli located in the right visual field produce activity within the region shown in the left hemisphere inferior view, along with activity in the medial and dorsal cortex (not shown); left visual field stimulation produces activity in the right hemisphere striate and extrastriate cortex (not shown). A = anterior; P = posterior; V = ventral; D = dorsal; M = medial; L = lateral; STG = superior temporal gyrus; MTG = middle temporal gyrus; ITS = inferior temporal sulcus; SF = Sylvian fissure; SPL = supramarginal gyrus; AG = angular gyrus; CeS = central sulcus; PCG = precentral gyrus; SFG = superior frontal gyrus; SFS = superior frontal sulcus; MFG = middle frontal gyrus; IFG = inferior frontal gyrus; CaS = calcarine sulcus; LG = lingual gyrus; PHG = parahippocampal gyrus; CoS = collateral sulcus; MeFG = medial fusiform gyrus; LaFG = lateral fusiform gyrus; PC = piriform cortex; OFG = orbitofrontal gyrus.
entially process a specific category of item rather than being specialized neural modules that process a single category. The fact that house processing was associated with activity in the parahippocampal cortex deserves further mention because, as noted above, this region has been considered part of the “where” pathway; as such, activity within this region may reflect visual spatial processing. In support of this possibility, the parahippocampal gyrus and collateral sulcus (extending into the medial fusiform gyrus) have been associated with preferential processing of stimuli that can be broadly described as visual context, including indoor or outdoor scenes, houses, furnished or empty rooms, and landscapes (Bar & Aminoff, 2003; Epstein & Kanwisher, 1998). In the present review, only two categories of particular relevance will be considered: faces (associated with activity in the fusiform gyrus) and visual context (associated with activity in the collateral sulcus extending into the medial fusiform gyrus and parahippocampal gyrus).

**Feature Specific**

In addition to describing visual items by their category, they can be described at a more detailed feature-specific level, including color or direction of motion. Research on monkeys has shown that specific regions are preferentially associated with processing each feature. For example, monkey extrastriate visual area V4 has been associated with color processing and is part of the ventral processing stream, whereas monkey visual areas MT and MST (middle temporal and medial superior temporal) have been associated with motion processing and are part of the dorsal processing stream (Livingstone & Hubel, 1988; Van Essen & Gallant, 1994). Similarly in humans, preferential color processing occurs in ventral extrastriate visual area V8 (the human homologue of the monkey color-processing region) that spans the collateral sulcus (Hadjikhani, Liu, Dale, Cavanagh, & Tootell, 1998; Zeki et al., 1991); whereas preferential motion processing occurs in MT+ (the human homologue of the monkey motion processing regions) that is typically located within the ascending limb of the inferior temporal sulcus (Huk, Dougherty, & Heeger, 2002; Tootell et al., 1995; Watson et al., 1993). An item’s precise spatial location in the visual field can also be considered a feature. By correlating visual field deficits with occipital lesions incurred during war (Holmes, 1945; Holmes & Lister, 1916; Inouye, 1909) or from more natural causes (Horton & Hoyt, 1991a, 1991b), visual field locations have been mapped onto the striate and extrastriate cortex. Such cortical retinotopic maps have been subsequently confirmed (and extended/refined) using fMRI (Engel, Glover, & Wandell, 1997; DeYoe et al., 1996; Sereno et al., 1995; Slotnick & Moo, 2003; Slotnick & Yantis, 2003). In addition to the point of fixation being mapped posteriorly in the occipital cortex with more eccentric locations being mapped more anteriorly, the retinotopic map is inverted and left-right reversed. That is, of most importance to the present review, the right visual field maps onto left hemisphere striate and extrastriate cortex, whereas the left visual field maps onto right hemisphere striate and extrastriate cortex (see Figure 2C).

**VISUAL ITEM MEMORY**

**Modality Specific**

The general format of the review will be to consider each visual memory study in turn, first highlighting the relevant results from each comparison of interest (followed by additional results from that comparison, for completeness). During the encoding phase of a visual item memory event-related fMRI study (Slotnick, Moo, Segal, & Hart, 2003), abstract shapes were presented on the left or right side of the display, and participants were instructed to remember each shape and its spatial location. During the item memory retrieval phase, encoded old shapes and new shapes were presented centrally, and participants responded as to whether each shape was “old” or “new.” Note that although memory for source/spatial location was also assessed in this study (and in other studies described below), discussion will be limited to those aspects of particular relevance (i.e., pertaining to visual item memory). To identify the neural regions associated with item memory retrieval, old shapes that were correctly remembered were contrasted with new shapes that were correctly rejected (this is currently the standard contrast in event-related designs to identify recognition memory-related activity). This contrast was associated with visual processing activity in the left lingual gyrus (BA18) and left parahippocampal gyrus (BA36) in addition to activity in the left middle temporal gyrus (BA21), precuneus (BA7), motor cortex (BA4), and right middle frontal gyrus (BA9). Such visual processing activity has been replicated and extended in a similar abstract visual shape item-memory fMRI study (Slotnick & Schacter, 2004) and echoes the results of two earlier PET studies conducted by Schacter and colleagues (1995, 1997), who were investigating memory for line drawings of simple geometrical objects. Because of the limited temporal resolution of PET, event-related analysis cannot be conducted with this method; rather, numerous instances of a given event type are presented in blocks, and the differential neural activity between blocks is assumed to reflect the cognitive component of interest. In these studies, recognition memory effects were obtained by contrasting blocks of old objects versus blocks of new objects, which is reasonable as the propor-
tion of “old” responses was greater during old object blocks as compared to new object blocks (and can be assumed to reflect a greater degree of item memory retrieval). In the initial study, object recognition memory was associated with visual processing activity in the fusiform and left parahippocampal gyri and the right hippocampal formation, with additional activity in the midbrain and left dorsolateral prefrontal cortex (BA10, BA44-46; Schacter et al., 1995). In the subsequent study, object recognition memory was associated with visual processing activity in bilateral parahippocampal gyri and bilateral hippocampal formations in addition to activity in the middle temporal gyrus (BA37; Schacter et al., 1997).

Cued recall paradigms have also been used to investigate modality-specific visual memory effects. In an fMRI experiment conducted by Wheeler, Petersen, and Buckner (2000), words were presented during the encoding phase (e.g., the word dog) along with either a detailed picture of the item (drawn by an artist) or the sound the item makes; participants were instructed to remember each picture or sound. During the retrieval phase, only encoded words were presented, and participants responded as to whether each word had previously been paired with a picture or a sound. The contrast of picture recall versus sound recall was associated with visual processing activity in the left fusiform gyrus (BA19), bilateral middle occipital gyri (BA19), and bilateral parietal cortex (BA7 and BA40 on the right) in addition to activity in the bilateral middle frontal gyrus (BA6). In a subsequent fMRI study using a similar paradigm (Wheeler & Buckner, 2003), this same contrast was associated with visual processing activity in the left fusiform gyrus (BA19/37), and bilateral parahippocampal gyr (BA20/36) in addition to activity in the bilateral cuneus (BA19/7) and medial and right frontal cortex (BA10/6). In a blocked fMRI design that did not explicitly employ cued recall, Vaidya, Zhao, Desmond, and Gabrieli (2002) presented participants with line drawings of common objects (pictures) or words during the encoding phase. During the retrieval phase, participants were presented with blocks of words, most of which had been previously presented as pictures or most of which had been previously presented as words. Participants responded as to whether each word was old or new, regardless of whether it had been previously presented as a picture or word; these words can be considered cues for recall of picture or word information. A contrast between picture blocks versus word blocks revealed a single locus of visual processing activity in the left fusiform gyrus (BA19, BA37).

The visual recognition and recall studies discussed above evoked item memory for abstract shapes, geometrical objects, or pictures. Visual item memory was associated with activity in both the ventral and dorsal visual processing streams. The most consistent activity, across studies, occurred in the fusiform gyrus (BA19/37) (Schacter et al., 1995; Vaidya et al., 2002; Wheeler & Buckner, 2003; Wheeler et al., 2000) and parahippocampal gyrus (Schacter et al., 1995, 1997; Slotnick, Moo, Segal, et al., 2003; Wheeler & Buckner, 2003). As visual item processing has been associated with these regions, these results suggest that visual item memory can be visual modality specific. This can be taken as evidence for a single dissociation given that visual item memory activated visual processing regions and did not generally activate neural regions associated with other modalities such as auditory processing regions. The evidence that immediately follows reveals item-memory-modality-specific double dissociations. Two of the studies described above that showed activity in visual processing regions during visual item memory also reported activity in auditory processing regions during memory for sounds (Wheeler & Buckner, 2003; Wheeler et al., 2000). In addition to other studies showing auditory word-processing activity during memory for sounds or spoken words (Nyberg, Habib, McIntosh, & Tulving, 2000; Schacter et al., 1996), memory for actions produces motor-processing region activity (Nyberg et al., 2001), and memory for odors produces olfactory-processing region activity (Gottfried, Smith, Rugg, & Dolan, 2004). These nonvisual modality-specific memory effects considered in combination with the visual recognition and recall effects in visual processing regions provide compelling evidence that visual item memory can be modality specific.

Findings from semantic memory, another form of long-term memory, also provide evidence for modality-specific memory effects. Items in semantic memory have already been encoded during the course of life (as it is defined here). As such, semantic memory paradigms have only a retrieval phase. In a PET study by Martin, Wiggs, Ungerleider, and Haxby (1996), participants were presented with line drawings of either animals or tools and silently named them (e.g., a bear or a saw; they also viewed visual noise patterns and nonsense objects). The animal versus tool contrast was associated with activity in striate cortex in addition to activity in the left frontal lobe, whereas the tool versus animal contrast was associated with activity in the premotor cortex (BA6) extending into the left lateral inferior frontal cortex (BA44) in addition to the right supramarginal gyrus, the left anterior cingulate (BA32), and the left middle temporal gyrus (this latter region will be discussed below). That animal naming produced preferential activity in striate cortex (presumably due to greater association with visual perceptual attributes) and tool naming pro-
duced preferential activity in the premotor cortex (presumably because tools are associated with motor action) can be taken to reflect modality-specific activity associated with semantic memory retrieval. In further support of motor modality-specific semantic memory effects, left premotor cortex (BA6) activity has also been observed in a blocked fMRI study that compared tools versus animals (using black-and-white pictures as stimuli; Chao & Martin, 2000) and a PET study that compared manipulable versus nonmanipulable objects (using color photographs as stimuli; Kellenbach, Brett, & Patterson, 2003).

**Domain Specific**

The studies described thus far show that both “what” and “where” pathways can be activated during visual item memory, paralleling the ventral and dorsal stream activity pattern observed during visual perception; however, they do not support domain specificity per se. Moscovitch, Kapur, Köhler, and Houle (1995) conducted a PET study to directly test for preferential “what” and “where” pathway activity during item memory. During the encoding phase, participants were sequentially presented with line drawings of three objects (e.g., a glass, an airplane, and a shirt) in pseudorandom spatial locations. During each trial of the retrieval blocks, a pair of displays was sequentially presented that included one display from the encoding phase in addition to the following: (a) for perceptual baseline blocks, another display from the encoding phase; (b) for object retrieval blocks, a display with objects in identical stimulus locations with one object that was new; and (c) for spatial retrieval blocks, a display with identical objects but with one differing in spatial location. In the perception blocks, participants responded as to whether the displays differed or were identical, whereas in the memory tasks, participants identified the altered display. Consistent with previous visual item memory results, both object memory blocks and spatial memory blocks, compared to baseline, were associated with activity in both ventral and dorsal visual processing streams. Critically, the object memory versus spatial memory contrast was associated with a single locus of ventral processing stream activity in the right inferior parietal/fusiform gyrus (BA37), whereas the spatial memory versus object memory contrast was associated with a single locus of dorsal processing stream activity in the inferior parietal lobule (BA40). These results show that visual memory for item identity and item spatial position can be associated with preferential “what” and “where” pathway activity, respectively, providing evidence for one type of domain specificity observed with visual perception processing.

Visual memory for item category constitutes another type of domain-specific evidence. In two fMRI experiments, Katanoda, Yoshikawa, and Sugishita (2000) asked participants to remember a series of novel faces during the encoding phase. Following a delay, the retrieval phase of one experiment consisted only of encoded faces (all-target) whereas the retrieval phase of the other experiment consisted of half encoded faces and half new faces (half-target). Of relevance is the contrast between the all-target versus half-target experiments, which assumes that a relatively greater degree of face memory retrieval occurred during the all-target experiment. This contrast was associated with activity only in bilateral fusiform gyri (BA19) and may be taken as evidence for domain specificity given that face processing has been associated with activity in this region. Complementing these findings, Burgess, Maguire, Spiers, and O’Keefe (2001) used event-related fMRI to explore the neural substrates associated with memory retrieval of visual context. During the encoding phase, participants explored a virtual reality town (like playing a video game) and received different objects (e.g., a teapot) from different people (e.g., a man in a vest; person context) in different places (e.g., a room with a bookshelf; visual context), and they were instructed to remember each object and its context. During the retrieval phase, participants were presented with two objects in the context of a previously seen person and place. Participants were cued to either identify which object was old, make a perceptual judgment comparing the two objects, identify which object was from that person, or identify which object was from that place. To isolate the process of memory for visual context from more general contextual memory per se, memory for places was contrasted with memory for people (only correct responses were analyzed). This contrast was associated with activity in bilateral parahippocampal gyrus (BA27/28, BA35), the region that has been associated with visual context, in addition to dorsal visual processing regions including right superior occipital gyrus (BA19) and the right parietal cortex (BA39/40) along with bilateral precuneus (BA7), bilateral parietal occipital sulci/retrosplenial cortex (BA18/17/23/29/30/31), cingulate cortex (BA23), and bilateral anterior prefrontal cortex (BA10/32). In another fMRI study (O’Craven and Kanwisher, 2000), black-and-white photographs of famous faces or campus buildings (both of which had already been familiar to the participants) were presented, and participants were instructed to silently name each one. In the retrieval phase, participants heard the names of encoded faces or buildings and were instructed to recall the corresponding previously encoded photographs. A contrast between face recall and place recall was associated with activity in the right fusiform gyrus, whereas the contrast between building recall and face recall was associated with activity in bilateral parahippocampal gyrus in addition to bilateral primary visual cortex. These results add further support to
the category-specific memory effects described above, reproducing the known domain specificity in perceptual processing where fusiform activity is associated with processing faces and parahippocampal activity is associated with processing visual context.

**Feature Specific**

The following studies manifested feature-specific effects, the first of which relate to memory for spatial location, reflecting perceptual retinotopic cortical organization (where left hemisphere striate and extrastriate cortex is associated with right visual field stimulation and right striate and extrastriate cortex is associated with left visual field stimulation; Figure 2C). In the encoding phase of an ERP study by Gratton, Corballis, and Jain (1997), participants were presented with simple symmetrical line patterns to the left and right of a fixation cross and responded as to whether each pattern was horizontally or vertically symmetrical. During the retrieval phase, old patterns and new patterns were presented at the center, and participants decided whether each shape was old or new. Correctly remembered old patterns that had been previously presented in the left hemifield as compared to those that had been previously presented in the right hemifield showed a differential activity profile on the right scalp and the opposite activity profile on the left scalp. A similar ERP finding was observed during memory for words that had been lateralized to the left or right hemifield during encoding (Fabiani, Stadler, & Wessels, 2000). These ERP memory effects are suggestive of activity in lateralized striate and extrastriate cortex (see Clark, Fan, & Hillyard, 1995; Di Russo, Martínez, Sereno, Pitzalis, & Hillyard, 2001); however, a method with high spatial resolution is necessary to obtain definitive results. In a recent event-related fMRI study (Slotnick & Schacter, 2005a), participants were presented with abstract shapes (filled with colored oriented lines) to the left and right of fixation during the encoding phase. They were instructed to remember each shape and its spatial location. During the retrieval phase, old shapes and new shapes were presented at fixation, and participants responded as to whether each shape was “old and on the left,” “old and on the right,” or “new.” Retrieval-related activity in the right fusiform gyrus (BA18) was associated with shapes previously presented on the left as compared to shapes previously presented on the right, whereas activity in the left lingual gyrus (BA18) was associated with shapes previously presented on the right as compared to shapes previously presented on the left. These hemisphere-lateralized memory effects provide direct feature-specific evidence that parallels the known retinotopic cortical organization associated with visual perception.

Martin et al.’s (1996) semantic memory study (described above) provides additional feature-specific evidence, in that tool naming, to a greater degree than animal naming, was associated with activity in the middle temporal gyrus, given that this region (which is located anterior and inferior to motion processing region MT+) has since been associated with tool motion (Beauchamp, Lee, Haxby, & Martin, 2002, 2003). These results were replicated and extended in a subsequent fMRI study that found greater activity associated with tools than animals in the middle temporal gyrus during viewing, matching, naming grayscale photographs, or reading words (Chao, Haxby, & Martin, 1999). It appears, then, that tool processing is associated with recalling the feature of tool motion (akin to perceiving the tool in motion).

**VISUAL WORKING MEMORY**

**Modality Specific**

An early PET study explored working memory for visual spatial location (Jonides et al., 1993). During each encoding phase of working memory blocks, three dots, in the peripheral visual field located on an imaginary circle centered on the fixation cross, were briefly presented. During the delay phase, the dots disappeared for 3 seconds (at which time participants were presumably holding the visual image of the dots in mind). During the retrieval phase, a small circle appeared that had an even chance of surrounding one of the previously presented dots or surrounding a previously blank region (on that trial); participants responded as to whether the circle surrounded one of the previously presented dots. The perception blocks, which served as a perceptual control, consisted of a fixation cross at the time of what would have been the previous encoding phase, a 3-second delay period, followed by a brief presentation of the fixation cross and 3 dots (identical to the working memory encoding phase), and then the addition of the probe circle; participants made the same decision. A contrast between working memory and perception was designed to isolate working-memory-related activity. This contrast was associated with visual processing activity in the extrastriate cortex (BA19) and parietal cortex (BA40) in addition to the premotor cortex (BA6) and prefrontal cortex (BA47). This visual processing activity is indicative of a modality-specific effect during visual working memory.

In a subsequent PET experiment conducted by the same group of investigators (Smith, Jonides, & Koepppe, 1996), the identical protocol was used to study visual spatial working memory and a similar protocol was used to study verbal working memory. The verbal working memory paradigm was the same except letters (instead of dots) were used as stimuli and participants were encour-
aged to use verbal codes. Visual working memory was associated with the same four regions of activity described above (BA19, BA40, BA6, and BA47). Of relevance, verbal working memory was associated with word-production-related activity in the inferior prefrontal cortex (BA44), a region that has been associated with verbal working memory in a number of studies (for a review, see Smith & Jonides, 1999). Considered together, the visual spatial working memory and verbal working memory results can be taken as evidence that working memory can be modality specific.

**Domain Specific**

Domain-specific visual working memory evidence also exists, primarily with regard to “what” and “where” visual spatial processing within ventral and dorsal prefrontal cortex (where ventral prefrontal cortex consists of the inferior and middle frontal gyri and dorsal prefrontal cortex consists of the superior frontal gyrus/sulcus; Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998). In a PET study by Courtney, Ungerleider, Keil, and Haxby (1996), participants viewed three photographs of faces sequentially at 1 of 24 display locations during the encoding phase. After a brief (500 ms) blank delay phase, the retrieval phase stimulus consisted of a face presented at one of the locations. Before each block, participants were instructed to remember either the three faces or the three spatial locations; during the retrieval phase, participants responded as to whether the face or spatial location (depending on the preceding instructional set) was the same as one of the three that were encoded. The contrast between working memory for faces (i.e., “what” processing) versus working memory for spatial location (i.e., “where” processing) was associated with ventral prefrontal cortex activity in the right inferior frontal cortex (BA9/45/46) in addition to activity in bilateral fusiform gyrus (BA18), right lingual gyrus (BA37), right parahippocampal gyrus (BA20/36), bilateral cerebellum, right thalamus, and right orbital frontal cortex (BA11). The opposite contrast between working memory for spatial location and faces was associated with dorsal prefrontal and parietal cortex activity in bilateral superior frontal sulcus (BA6/8) and bilateral inferior and superior parietal cortex (BA40 and BA7) in addition to activity in the right striate cortex (BA17) and left middle occipital cortex (BA19). These results are in line with a domain-specific prefrontal cortex processing view of working memory, as the face greater than spatial location contrast was associated with inferior prefrontal cortex activity whereas the spatial location greater than face contrast was associated with dorsal prefrontal and parietal cortex activity. A similar paradigm was subsequently conducted with event-related fMRI (Courtney et al., 1998), with the main difference being a longer working memory delay phase (9 seconds); this technique has the advantage that activity associated with the working memory delay phase can be isolated from the encoding and retrieval phases. Replicating the previous results, there was greater delay period activity in the left inferior frontal cortex (and to a lesser degree in the left middle frontal cortex) during working memory for faces than during working memory for spatial locations, and there was greater delay period activity in bilateral superior frontal sulcus during working memory for spatial locations than during working memory for faces. Furthermore, the spatial-working-memory-specific activity in the superior frontal sulcus was just anterior to activity associated with eye movements, a finding used to argue that this region may be specialized for spatial working memory function (a point that will be considered further below). A more recent event-related fMRI experiment using a similar paradigm (also with a 9-second delay phase) was conducted by Sala, Rămă, and Courtney (2003) and extended these findings by using grayscale photographs of houses as stimuli. In particular, working memory for houses versus working memory for spatial location was associated with ventral prefrontal activity in bilateral inferior frontal gyrus/insula/middle frontal gyrus in addition to activity in the left fusiform gyrus and cingulate cortex/presupplementary motor area (note that for this contrast, houses might be better construed as to-be-remembered items rather than visual context), whereas working memory for spatial location versus working memory for houses was associated with dorsal prefrontal and parietal activity in bilateral superior frontal sulcus and bilateral intraparietal sulcus. A similar pattern of fMRI results has also been reported in a study by Belger et al. (1998), in which working memory for abstract shapes was predominantly associated with activity in the inferior frontal gyrus and middle frontal gyrus along with mostly temporal cortex activity, whereas working memory for the spatial location of white squares was predominantly associated with activity in the superior frontal gyrus along with mostly parietal activity (direct contrasts between working memory for abstract shapes and spatial locations were not conducted). This pattern of results, across a range of stimulus materials, provides compelling evidence that ventral prefrontal cortex is associated with working memory for object identity whereas the dorsal prefrontal cortex is associated with working memory for spatial location.

However, whether object working memory and spatial working memory are associated with ventral and dorsal prefrontal cortex is a matter of ongoing debate. Some investigators have not reported differential prefrontal cortex activity during working memory for item identity as compared to working memory for item spatial location (i.e., null findings; D’Esposito et al., 1998; Owen et
al., 1998; Postle & D’Esposito, 1999) and rather have found evidence that ventral and dorsal prefrontal cortex are associated with maintenance and manipulation of items in working memory, respectively (D’Esposito, Postle, Ballard, & Lease, 1999; Owen, Evans, & Petrides, 1996; Postle, Berger, & D’Esposito, 1999; Postle, Berger, Taich, & D’Esposito, 2000). It should be noted that the object-spatial and maintenance-manipulation views of prefrontal cortex function are in no way incompatible, and there is compelling evidence that both views are correct (see meta-analyses by Smith & Jonides, 1999, and D’Esposito et al., 1998). Still, the debate continues. In an event-related fMRI study by Postle et al. (2000), equivalent spatial-working-memory-delay-related activity and eye-movement-related activity was reported in the superior frontal sulcus (i.e., a null finding that failed to replicate Courtney et al., 1998). If correct, this observation would undermine the evidence for spatial working memory specialization in dorsal prefrontal cortex (and the object-spatial/ventral-dorsal prefrontal cortex working memory distinction more generally). However, in a reanalysis of Postle et al.’s (2000) data, event-related time course analysis (rather than the more conventional beta-weight analysis) revealed greater spatial-working-memory-related activity than eye-movement-related activity in the superior frontal sulcus (Slotnick, in press). Thus, the failure of some studies to report differential object-spatial-working-memory-related prefrontal cortex activity might be attributable to a lack of sensitivity. Indeed, studies that have reported significant differential activity in the prefrontal cortex have consistently found that ventral prefrontal cortex is associated with object working memory and that dorsal prefrontal cortex is associated with spatial working memory.

There is also category-related evidence that working memory is domain specific. In an event-related fMRI study conducted by Ranganath, DeGutis, and D’Esposito (2004), participants viewed grayscale pictures of faces and houses (three total) during the encoding phase and were instructed to remember either the face(s) or the house(s) during the delay phase (which lasted 12 seconds). During the retrieval phase, participants were presented with an old or new face on face working memory trials or an old or new house on house working memory trials and responded as to whether the item was old or new. Working memory delay effects were assessed within fusiform regions specialized for face processing and parahippocampal regions specialized for scene processing, identified in separate fMRI scans by contrasting face perception versus scene perception or vice versa. Within the fusiform regions, the face working memory delay period was associated with an increase in activity that was significantly greater than that associated with the scene working memory delay period. In the parahippocampal regions, the opposite pattern of activity was observed, with preferential delay phase activity associated with working memory for scenes. Similar double dissociations between fusiform cortex and more medial ventral temporal cortex have been observed in other studies comparing face and house working memory delay period activity (Ranganath, Cohen, Dam, & D’Esposito, 2004; Sala et al., 2003). These results indicate there can be domain-specific activity that reflects the known category-related cortical processing architecture associated with visual perception.

**Feature Specific**

As with visual item memory, visual working memory feature-specific effects consist of retinotopically organized visual area activity associated with memory for spatial location. In the encoding phase of an fMRI study by Awh et al. (1999), participants were sequentially presented with three false-font characters at a pseudo-random location within either the right or left hemifield. During the (6-second) delay phase, a flashing checkerboard display was presented in both hemifields to evoke a response in striate and extrastriate cortex. During the retrieval phase, a probe was presented, and participants responded as to whether it occupied one of the encoded locations. Working memory for stimuli in the right or left visual field was associated with activity in the contralateral occipital cortex. This spatial working memory contralateral effect has subsequently been replicated, being attributed to spatial selective attention (Awh, Anllo-Vento, & Hillyard, 2000; Postle, Awh, Jonides, Smith, & D’Esposito, 2004). Although these effects may be due to selective attention toward the stimuli locations during the working memory delay period, they may also be attributed to visual imagery of the encoded stimulus display given that imagery effects can be retinotopic (e.g., Kosslyn et al., 1993; Slotnick, Thompson, & Kosslyn, in press). Regardless of the underlying cognitive processes giving rise to these spatial working memory effects, they are entirely consistent with the known retinotopic organization associated with visual perception.

**DISCUSSION**

The evidence considered in the present review provides compelling support that visual item memory (including semantic memory) and visual working memory are associated with modality-specific, domain-specific, and feature-specific regions that have been associated with visual perception (see Figure 3).
The fact that visual memory can be associated with modality-specific, domain-specific, and feature-specific cortical activity might invite the interpretation that such activity is critical for conscious (explicit) memory. However, without further assessment, such an interpretation would be premature (see Caramazza & Shelton, 1998).

In an item memory event-related fMRI study that employed abstract visual shapes (filled with colored oriented lines), we determined whether occipitotemporal activity (BA17, BA18, BA19, BA37) was associated with conscious or nonconscious memory (Slotnick & Schacter, 2004). Specifically, old-hit– versus old-miss– related activity was assumed to be associated with conscious memory (i.e., remembering vs. forgetting, in which neural activity tracks behavioral response; Wheeler & Buckner, 2003), whereas old-hit and old-miss activity (relative to new-correction rejection activity) was assumed to be associated with nonconscious memory (i.e., neural activity is independent of behavioral response; Rugg et al., 1998). We found that activity in later visual processing regions (BA19, BA37) was associated with conscious memory, suggesting that previously reported activity in these regions (Vaidya et al., 2002; Wheeler & Buckner, 2003; Wheeler et al., 2000) was associated with conscious processing. Furthermore, activity in earlier visual regions (BA17, BA18) was associated with nonconscious memory. We argued that this earlier visual region activity might reflect repetition priming (a type of implicit memory), as repetition-related increases in activity—as opposed to typically observed repetition-priming-related decreases in activity (Buckner et al., 1995, 1998; Schacter & Buckner, 1998; Wiggs & Martin, 1998)—have been reported when unfamiliar faces or novel objects were used as stimuli (Schacter et al., 1995; Uecker et al., 1997; Henson, Shallice, & Dolan, 2000).

The feature-specific item memory retinotopic effects of Slotnick and Schacter (2005a) provide further evidence that earlier ventral occipital activity can reflect nonconscious memory. As described above, memory for hemifield lateralized abstract shapes was associated with activity in contralateral extrastriate cortex (BA18). As before, this might be assumed to be associated with conscious memory for each shape lateralized to the appropriate hemifield. However, a subsequent analysis revealed that this contralateral feature-specific memory effect did not depend, to a significant degree, on response accuracy. That is, the same contralateral effect was observed whether the visual shape and its spatial
location were remembered or forgotten, which argues against this effect being associated with conscious memory. Still, this is a null finding and must be treated with caution. It may well be the case that there is activity associated with conscious memory in some instances, but fMRI may not have sufficient sensitivity or is not measuring the appropriate neural signal (such as the frequency of the neural response, as is discussed below).

These results indicate that one cannot assume visual-memory-related activity is associated with conscious memory simply because an explicit memory task was employed (nor should it be assumed that an implicit memory task will evoke activity associated only with nonconscious memory). Furthermore, there is evidence that neural activity may also reflect different levels of conscious awareness (rather than simply conscious or nonconscious processing). For instance, conscious memory can be associated with the subjective experience of “remembering,” memory that includes contextual details of a previous event, or “knowing,” memory without contextual details (Tulving, 1985), and remembering consistently evokes greater activity than knowing in the medial temporal lobe (hippocampus and parahippocampal gyrus) and visual processing regions (for review, see Slotnick & Schacter, 2005b). It is hoped that the previous examples might spur memory researchers to test whether memory-related activity reflects nonconscious or conscious processing (or even assess the degree of conscious processing), rather than assuming such activity reflects conscious memory.

**Visual Memory Binding**

With regard to the activity in disparate neural regions that does reflect conscious memory of visual items, how might that information come together to construct a unified memory? This question illustrates the binding problem and is central to understanding the mechanism of visual memory construction. Cortical electrode recording in cats and monkeys indicates that distinct neural regions can synchronously oscillate in the gamma frequency range (20-60 Hz) which may serve to bind the information associated with each region into a visual perceptor (Fries, Roelfsema, Engel, Konig, & Singer, 1997; Gray, Konig, Engel, & Singer, 1989; Gray & Singer, 1989; Kreiter & Singer, 1996; Singer & Gray, 1995). Furthermore, simultaneous thalamic and depth electrode recording in cats indicates that gamma cortical rhythms are synchronized with neural activity in the thalamus (Amzica, Neckelmann, & Steriade, 1997; Steriade, 1997; Steriade, Contreras, Amzica, & Timofeev, 1996; Steriade, Dossi, Pare, & Oakson, 1991). In humans, there is complementary evidence that gamma activity occurs during visual memory. In the delay period of a visual working memory paradigm that used abstract shapes as stimuli, gamma activity on occipitotemporal scalp has been reported using electroencephalography (Tallon-Baudry, Bertrand, Peronnet, & Pernier, 1998). During semantic memory recall of visual objects, we have reported synchronized gamma activity in the thalamus and occipital scalp, using combined thalamic depth recording and electroencephalography (Slotnick, Moo, Kraut, Lesser, & Hart, 2003). The visual-memory-related gamma activity observed in these studies may reflect binding of visual features or components to create a unified memory; we have hypothesized the underlying mechanism depends, to some degree, on corticothalamic-cortical circuitry (see Slotnick, Moo, Kraut, et al., 2003). In addition to gamma frequency results, analysis within other frequency bands will most assuredly shed additional light on the mechanisms underlying visual memory (e.g., 4-8 Hz theta activity; see Kahana, Seelig, & Madsen, 2001). These results underscore the possibility that the mechanisms of visual memory may well benefit from our knowledge of the mechanisms underlying visual perception.

It has also been argued that the hippocampus is necessary to bind information from distributed cortical regions during item memory, although the hippocampus is not necessary for working memory (for a detailed review, see Squire, 1992). The present review has attempted to bridge visual item memory retrieval and the visual working memory delay period, in that both are associated with a unified memory that is constructed from features or components from disparate cortical regions. As such, the fact that item memory retrieval requires the hippocampus whereas working memory does not suggests the hippocampus is not associated with visual memory binding per se. Rather, the hippocampus may serve a function that specifically relates to long-term memory (e.g., forming connections via long-term potentiation; Malenka & Nicoll, 1999) that may precede visual memory binding. Future studies will be required to delineate the precise roles of the thalamus and hippocampus in visual memory binding.

**CONCLUSION**

The present review has shown that human visual memory can evoke activity in modality-specific, domain-specific, and feature-specific processing regions associated with visual perception. Although this visual perception/visual memory framework should be considered only a working model, it is hoped that it may provide a guide for future research. Furthermore, considering the degree of common neural substrates associated with visual memory and visual perception, the mechanisms underlying visual memory are expected to benefit from research into the mechanisms underlying visual
perception. For instance, both visual percepts and visual memories (and memories more generally) have been construed as constructed from features or components from disparate cortical regions. Recruitment of the same or similar neural mechanisms during the percept or memory construction process would be the most straightforward solution to the common binding problem.

REFERENCES


