CHAPTER 28

The Cognitive Neuroscience of Memory and Consciousness

Scott D. Slotnick and Daniel L. Schacter

Abstract

In this chapter, we delineate the neural activity associated with conscious memories characterized by different degrees of ‘retrieval content’ (i.e. sensory/contextual detail). Based primarily on neuroimaging evidence, we identify the neural regions that are associated most consistently with the following conscious memory processes: retrieval success versus retrieval attempt, remembering versus knowing, and true recognition versus false recognition. A number of patterns emerge from the comparison of memories with high retrieval content (i.e., retrieval success, remembering, and true recognition) and memories with low retrieval content (i.e., retrieval attempt, knowing, and false recognition). Memories with both high and low retrieval content are associated with activity in the prefrontal and parietal cortex, indicating that these regions are generally associated with retrieval. There is also evidence that memories with low retrieval content are associated with activity in the prefrontal cortex to a greater degree than memories with high retrieval content, suggesting that low retrieval content memories are associated with greater post-retrieval monitoring (although this activity does not necessarily reflect differential retrieval content per se). Finally, memories with high retrieval content, to a greater degree than memories with low retrieval content, are associated with activity in the parietal cortex and sensory cortex (along with the medial temporal lobe for retrieval success > attempt and remembering > knowing). This increased activity in sensory cortex (and medial temporal lobe) for memories with high retrieval content indicates that conscious memories are constructed by reactivation of encoded item features at retrieval.

Introduction

In 1985, Endel Tulving lamented the lack of interest in the topic of memory and consciousness shown by past and present memory researchers: "One can read article after article on memory, or consult book after book, without encountering the term
'consciousness.' Such a state of affairs must be regarded as rather curious. One might think that memory should have something to do with remembering, and remembering is a conscious experience (Tulving, 1985b, p. 11).“

Though Tulving provided an accurate assessment of the field at the time, the year in which he voiced his complaint proved to be a kind of turning point in research on memory and consciousness. Tulving’s (1985b) own article focused on the important distinction between remembering, which involves specific recollections of past experiences, and knowing, which involves a general sense of familiarity without specific recollection, and introduced seminal techniques for experimentally assessing these two forms of memory. In a different paper published that same year, Tulving (1985a) argued that each of three dissociable memory systems is uniquely associated with a particular type of consciousness. Specifically, he contended that procedural memory (learning of motor, perceptual, and cognitive skills) is associated with anoetic or “nonknowing” consciousness, which entails simple awareness of external stimuli; semantic memory (general factual knowledge) is associated with noetic or “knowing” consciousness; and episodic memory (recollection of personal experiences) is associated with autonoetic or “self-knowing” consciousness.

Finally, in 1985 Graf and Schacter introduced the related distinction between implicit and explicit memory. According to Graf and Schacter (1985), explicit memory refers to the conscious recollection of previous experiences, as revealed by standard tests of recall and recognition that require intentional retrieval of previously acquired information. Implicit memory, by contrast, refers to non-conscious effects of past experiences on subsequent behavior and performance, such as priming or skill learning, that are revealed by tests that do not require conscious recollection of previous experiences (for precursors, see also Cermák, 1982; Moscovitch, 1984). During the 20 years that have elapsed since the publication of these papers, a vast amount of research has been published on the distinctions that they introduced. For example, many cognitive studies have used the techniques introduced by Tulving (1985b) to delineate the functional and phenomenological characteristics of remembering and knowing (for reviews and contrasting perspectives, see Dunn, 2004; Gardiner, Ramponi, & Richardson-Klavehn, 2002). Likewise, cognitive studies have also explored numerous aspects of the relation between implicit and explicit forms of memory (for reviews, see Roediger & McDermott, 1993; Schacter, 1987; Schacter & Curran, 2000).

Although purely cognitive studies have played a significant role in advancing our understanding of memory and consciousness, cognitive neuroscience studies – which attempt to elucidate the nature of, and relations between, the brain systems and processes that support various forms of memory – have also been critically important. Indeed, much of the impetus for the distinction between implicit and explicit memory was provided initially by neuropsychological studies of amnesic patients, who exhibit severe impairment of explicit memory for previous experiences as a result of damage to the hippocampus and related structures in the medial temporal lobe (MTL; e.g., Moscovitch, Vriezen, & Goshen-Gottstein, 1993; Nadel & Moscovitch, 1997; Squire, 1992; Squire, Stark, & Clark, 2004). Nonetheless, it has been demonstrated repeatedly that conditions exist in which amnesics can exhibit robust and sometimes normal implicit memory for aspects of prior experiences as exemplified by such phenomena as preserved priming and skill learning (for recent reviews, see Gooding, Mayes, & van Eijk, 2000; Schacter, Dobbins & Schnyer, 2004). Studies of other neuropsychological syndromes have likewise revealed dissociations between implicit and explicit forms of perception, language, and related cognitive and motor processes (e.g., Goodale & Westwood,
Although neuropsychological studies have been crucial to advancing our understanding of the relation between memory and consciousness, during the past decade cognitive neuroscience analyses have focused increasingly on research using functional neuroimaging techniques, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). A vast amount of neuroimaging research has been published, and much of it is well beyond the scope of this chapter. However, we believe that several lines of research concerned with elucidating the neural correlates of explicit memory processes do provide useful insights into the cognitive neuroscience of memory and consciousness (for reviews of neuroimaging the cognitive neuroscience of memory and consciousness, see Henson, 2003; Schacter, McAndrews, & Moscovitch, 1998; Warrington & Weiskrantz, 1974; Young, 1994).

Many neuroimaging studies of explicit memory have used a recognition paradigm, where items such as words or objects are studied, and then on a subsequent test, these old items are randomly intermixed with new items, and participants decide whether each item is “old” or “new.” Item recognition has been associated most consistently with activity in three neural regions: (1) prefrontal cortex (anterior and dorsolateral), (2) parietal cortex, and (3) the MTL (for reviews, see Buckner & Schacter, 2004; Buckner & Wheeler, 2001; Slotnick, Moo, Segal, & Hart, 2003; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994). The functional role(s) subserved by each of these regions is currently an active area of investigation, as we discuss later in this chapter. At a very general level, prefrontal cortex has been associated with the control of retrieval (e.g., increases in activity that correlate with retrieval demands; Velanova et al., 2003; Wheeler & Buckner, 2003) and in addition has been associated with post-retrieval monitoring (for a review, see Schacter & Slotnick, 2004). Parietal cortex, particularly in Brodmann Area (BA) 39/40, has recently been associated with the tendency to make “old” responses (Velanova et al., 2003; Wheeler & Buckner, 2003). The MTL, as mentioned previously, is necessary for explicit memory, with the hippocampus proper possibly serving the role of binding together information from disparate cortical regions (Squire, 1992). That is, the hippocampus may serve a central role in combining disparate features to construct a unitary memory (e.g., Moscovitch, 1994; Schacter, Norman, & Koutstaal, 1998a; Squire, 1992).

Providing support for this constructive view of memory, recent neuroimaging evidence indicates that explicit memory evokes activity in the appropriate domain-specific processing regions (i.e., retrieval-related reactivation of processing regions associated with memorial encoding). Specifically, memory for actions activates motor processing regions (Nyberg, Petersson, Nilsson, Sandblom, Åberg, & Ingvar, 2001), memory for sounds activates auditory processing regions (Nyberg, Habib, McIntosh, & Tulving, 2000; Wheeler & Buckner, 2000; Wheeler, Petersen, & Buckner, 2003), memory for odors activates olfactory processing regions (Gottfried, Smith, Rugg, & Dolan, 2004), and memory for visual stimuli (e.g., shapes or objects) activates occipital-temporal regions in the ventral visual processing stream (Moscovitch, Kapur, Köhler, & Houle, 1995; Slotnick et al., 2003; Vaidya, Zhao, Desmond, & Gabrieli, 2002; Wheeler & Buckner, 2003; Wheeler et al., 2000). Such domain-specific sensory reactivation is typically taken as evidence for the conscious re-experiencing of sensory attributes of items from the study episode.

This chapter considers three lines of research that have examined aspects of this memory-related sensory/contextual activity and the associated subjective experience (or phenomenal consciousness; see Block, 1995; also referred to as the ‘contents of consciousness’ or more simply ‘retrieval
content'; Wheeler & Buckner, 2003). We believe that each of these lines of research has provided new information regarding the neural underpinnings of conscious experiences of remembering. First, we consider attempts to separate explicit retrieval into separate components that can be grouped broadly into two categories: retrieval success, which involves the recovery of information presented during a prior study episode, and retrieval attempt, which refers to strategic processes involved in explicit retrieval that operate even when recovery is not successful. Evidence from recent neuroimaging studies points toward different neural substrates subserving these two broad classes of conscious memory processes. Second, we discuss imaging experiments concerned with the distinction between remembering and knowing (Tulving, 1985b) that examine how neural activity correlates with differing degrees or types of conscious experiences. Third, we consider recent work concerned with delineating the neural substrates of true versus false memories, where the role of sensory reactivation in the conscious experience of remembering has been examined in the context of questions concerning the accuracy of explicit retrieval. Although we focus on neuroimaging studies in each of the three lines of research, we also discuss, when relevant, complementary data from neuropsychological studies of brain-damaged patients.

Neural Substrates of Retrieval Success Versus Attempt

When a brain region shows changes in activity during explicit retrieval, the changes are not necessarily associated with the conscious experience of successfully recovering previously studied information. Such changes could instead reflect, entirely or in part, conscious processes involved in the deployment of attention or effort when individuals attempt to retrieve the target material, independent of whether retrieval is successful. Once neuroimaging studies of episodic memory had demonstrated that explicit retrieval is accompanied by activation in specific brain regions – most prominently, regions within prefrontal cortex, but also within the MTL (e.g., Schacter, Alpert, Savage, Rauch, & Albert, 1996a; Squire et al., 1992; Tulving et al., 1994) – it became important to specify further the nature of the observed activity. Early PET studies adopted two main experimental approaches to this issue: (1) producing high and low levels of successful retrieval by manipulating study conditions and (2) manipulating the number of previously studied items that appear during a particular test. We briefly summarize studies that have used each type of approach.

In a PET study by Schacter et al. (1996a), subjects studied some words four times and judged the number of meanings associated with each item (high-recall condition); they studied other words once and judged the number of t-junctions in each item (low-recall condition). Subjects were then scanned during an explicit retrieval task (stem-cued recall, e.g., tab for table), with separate scans for high-recall words and low-recall words. The logic underlying the experiment is that regions that are selectively activated during the high-recall condition, when subjects correctly recall a large proportion of the study list words, are preferentially associated with successful conscious recollection; by contrast, regions that are activated during the low-recall condition, when subjects retrieve only a few study lists words, are preferentially associated with retrieval attempt. Analysis of PET data revealed blood flow increases in the hippocampal formation during the high-recall but not the low-recall condition, and a significant difference between the two conditions, thereby suggesting that hippocampal activation is associated with some aspect of the successful conscious recall of a previously studied word, rather than retrieval attempt (see also, Nyberg, McIntosh, Houle, Nilsson, & Tulving, 1996; Rugg, Fletcher, Frith, Frackowiak, & Dolan, 1997). Schacter et al. (1996a) also found that anterior/dorsolateral areas within prefrontal cortex were preferentially activated in the
low-recall condition, thus raising the possibility that blood flow increases in anterior prefrontal cortex during stem-cued recall are associated with retrieval orientation effects (cf., Nyberg et al., 1995). Such findings accord well with theoretical proposals that have linked the MTL/hippocampal region with the automatic recovery of stored information and regions within prefrontal cortex with strategic aspects of retrieval (e.g., Moscovitch, 1994).

In a related PET study by Rugg et al. (1997), subjects studied word lists and either generated sentences for each word (deep encoding) or made judgments about the letters in each word (shallow encoding). Following each type of encoding task, they were given either an old-new recognition test (intentional retrieval) or an animate/inanimate decision task (unintentional retrieval). Deep encoding produced more accurate memory on the intentional retrieval task. Performance was at ceiling levels on the unintentional task, but subjects reported spontaneously noticing that test words came from the study list more often after deep than shallow encoding, perhaps providing a rough index of unintentional conscious recollection. There was greater activation in left MTL areas after deep encoding than after shallow encoding during both intentional and unintentional retrieval. Thus, these data suggest that hippocampal activity during retrieval is observed with high levels of conscious recollection, regardless of whether subjects voluntarily try to remember the study list items. By contrast, there was greater right prefrontal activation during intentional retrieval than during unintentional retrieval after both deep and shallow encoding.

Several PET studies have attempted to separate retrieval success and retrieval attempt by manipulating the proportion of old items presented to subjects during a particular scan. The reasoning here is that presenting large numbers of old items during a particular scan will produce more successful retrieval than presenting only a few old items. In general, these studies focused on issues concerning the characterization of retrieval-related activation observed in right anterior prefrontal cortex. However, results from these studies were inconclusive, with some evidence linking right prefrontal activation with retrieval attempt (e.g., Kapur, Craik, Jones, Brown, Houle, & Tulving, 1995) and others reporting evidence for retrieval success effects (e.g., Rugg, Fletcher, Frith, Frackowiak, & Dolan, 1996; for an attempt to reconcile some of these conflicting early results, see Wagner, Desmond, Glover, & Gabrieli, 1998).

The development of event-related fMRI in the late 1990s provided a more direct means of examining brain activations associated with retrieval success and retrieval attempt. The PET studies reviewed above used blocked designs in which items from different conditions were presented in separate blocks, and data concerning brain activity were collapsed across subjects’ behavioral responses. Taking advantage of the superior temporal resolution of fMRI compared with PET, event-related fMRI allows intermixing of items from different conditions and, more importantly, permits analysis of brain activity conditional on subjects’ responses (Dale & Buckner, 1997). Thus, for example, in a recognition memory task, “old” and “new” responses can be analyzed separately for old and new items. Thus, retrieval success should be maximal when subjects make “old” responses to old items (hits) and minimized when subjects make “new” responses to new items (correct rejections).

A number of studies have used event-related fMRI to examine retrieval success versus retrieval attempt with an old-new recognition test for previously studied items intermixed with new, non-studied items. The critical comparison involves a contrast of brain activity during hit versus correct rejection trials (e.g., Buckner, Koutstaal, Schacter, Dale, Rotte, & Rosen, 1998; Schacter, Buckner, Koutstaal, Dale, & Rosen, 1997a). However, as discussed by Konishi, Wheeler, Donaldson, and Buckner (2000), these failures to observe
evidence for retrieval success effects likely reflected technical limitations of early event-related fMRI procedures, such as low statistical power resulting from the use of long intertrial intervals and a correspondingly low number of items per experimental condition. Consistent with this possibility, studies using more powerful event-related methods revealed evidence for greater activation during hits than during correct rejections in a number of cortical regions, most consistently in prefrontal and parietal cortices (e.g., Konishi et al., 2000; McDermott, Jones, Petersen, Lageman, & Rodeiger, 2000; Nolde, Johnson, & D’Esposito, 1998) but also in the MTL (as is discussed below).

The results of the foregoing studies are consistent with the conclusion that regions within prefrontal and parietal cortices are specifically related to successful conscious recollection of some aspects of a previous experience. However, this conclusion depends critically on the assumption that comparing hits with correct rejections isolates successful retrieval. Although the assumption appears straightforward enough, the comparison between hits and correct rejections necessarily confounds subjects’ responses (“old” or “new”) and item type (old and new). It is conceivable, therefore, that hit greater than correct rejection-related brain activations do not exclusively reflect differences in conscious experience related to subjects’ responses (“old” versus calling it “new”), but instead reflect differences in responses to old versus new items, irrespective of subjects’ experiences. For example, differential responses to old and new items might reflect the occurrence of priming or related processes that can occur independently of conscious memory (Schacter & Buckner, 1998; Schacter et al., 2004; Wiggs & Martin, 1998).

Dobbins, Rice, Wagner, and Schacter (2003) approached this issue within the context of the theoretical distinction between recollection (i.e., memory for the contextual details of a prior encounter) and familiarity (i.e., recognition without recollection of contextual details). Both recollection and familiarity can, in principle, operate on a particular memory test (for a contrasting view, see Slotnick & Dodson, 2005). Moreover, each of the two processes are potentially separable into the two components on which we have focused in this section of the chapter, retrieval success and retrieval attempt. Dobbins et al. used the closely related phrase, “retrieval orientation” to refer to the extent to which subjects recruit each process during particular retrieval tasks. With respect to the issues raised in the preceding paragraph, Dobbins et al. (2003) noted that, when presented with new items, subjects could rely entirely on familiarity-based processes, rejecting new items when they are not familiar, and might not even attempt to engage in recollection-based retrieval. Thus, it is conceivable that previous findings of prefrontal and parietal activations associated with hits greater than correct rejections might reflect attempted recollective retrieval, rather than successful conscious recollection.

To address this issue, Dobbins et al. (2003) used a different type of experimental design in which all items had been presented previously, and task demands were varied to require differential reliance on recollection and familiarity. Prior to scanning, subjects were presented visually with a long list of nouns, and then they alternated between two semantic encoding tasks (pleasant/unpleasant and concrete/abstract judgments). Subjects were then scanned during two different two-alternative forced-choice tests: a source memory test and a recency memory test. During source memory, subjects selected the member of the pair previously associated with a particular encoding task; that is, they had to recollect some type of detail associated with the particular encoding judgment performed earlier. In contrast, the recency judgment required subjects to select the most recently encountered item of the pair, regardless of how it had been encoded. The source memory test is assumed to rely on recollection, whereas recency decisions can rely on a familiarity signal. Furthermore, successful and unsuccessful trials within each retrieval task were contrasted to determine whether retrieval success effects occurred in
overlapping or dissimilar brain regions compared to those associated with each retrieval orientation.

Results revealed left lateral prefrontal and parietal activations that distinguished attempted source recollection from judgments of relative recency; these retrieval attempt or orientation effects were largely independent of retrieval success. Importantly, these activations occurred largely in the same left prefrontal and parietal regions that had been previously identified with retrieval success. Because these regions were not associated with successful retrieval in the Dobbins et al. (2003) design, which controlled for old-new item differences present in previous studies, it is plausible that the prefrontal and parietal activations in earlier studies reflect attempted, rather than successful, conscious recollection (for further relevant analyses, see Dobbins, Foley, Schacter, & Wagner, 2002). In contrast, Dobbins et al. (2003) found that MTL structures (hippocampus and parahippocampal gyrus) were differentially more active during successful recollection, showing similarly reduced responses during failed source recollection and judgments of recency. These findings complement previous data linking MTL regions with successful conscious recollection (e.g., Maril, Simons, Schwartz, Mitchell, & Schacter, 2003; Rugg et al., 1997; Schacter et al., 1996a; but see, Buckner et al., 1998; Rugg, Henson, & Robb, 2003), as well as other results from related paradigms considered later in the chapter.

Kahn, Davachi, and Wagner (2004) have provided converging evidence on the foregoing conclusions using an old-new recognition test for previously presented words and new words in which subjects also made a source memory judgment (whether they had read a word at study or imagined a scene related to the word). They concluded that left prefrontal/parietal regions are related to attempted recollection of source information, but not to successful recollection of that information; by contrast, MTL activation (in the parahippocampal region) was related to successful source recollection. Importantly, Kahn et al. (2004) also provided evidence indicating that left frontal/parietal activity is related to familiarity-based retrieval success. Thus, the general distinction between retrieval success and retrieval attempt (or orientation) may be too coarse to prove useful theoretically. Instead, it may be necessary to specify a particular form of retrieval to make sense of neuroimaging data concerning the neural correlates of successful and attempted retrieval (e.g., recollection versus familiarity, or remembering versus knowing, which are considered in detail below).

The results summarized in this section indicate that neuroimaging studies are beginning to dissociate components of conscious retrieval that are related to activity in particular brain regions. In particular, three patterns of results can be observed (Table 28.1, left column). First, both retrieval success and retrieval attempt

<table>
<thead>
<tr>
<th>Region</th>
<th>Prefrontal Cortex</th>
<th>Parietal Cortex</th>
<th>MTL</th>
<th>Sensory Cortex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval Success &amp; Attempt</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retrieval Success &gt; Attempt</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retrieval Attempt &gt; Success</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remembering &amp; Knowing</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Remembering &gt; Knowing</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowing &gt; Remembering</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>True &amp; False Recognition</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>True &gt; False Recognition</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>False &gt; True Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Regions of common (&) and differential (>) activity were identified via review of the neuroimaging literature.*
were associated with activity in the prefrontal cortex and parietal cortex. Second, retrieval success to a greater degree than retrieval attempt was associated with activity in the MTL. Third, there is some evidence that retrieval attempt may be associated with greater prefrontal cortex activity than with retrieval success. As research in this area progresses, increasingly finer distinctions will be made regarding the neural substrates associated with particular aspects of conscious memorial experience. Of note, the fact that retrieval success – which can be assumed to reflect greater retrieval content than retrieval attempt – is preferentially associated with the MTL suggests this region plays a role during conscious remembering.

Neural Activity Associated with Remembering and Knowing

We reviewed research in the preceding section that attempts to dissociate recollection and familiarity by manipulating task demands. However, as noted earlier, recollection and familiarity can be assessed directly by asking participants about their subjective experiences during a memory task; that is, to classify “old” responses based on the associated memorial experience of remembering or knowing. Remember responses indicate recollection of specific contextual detail associated with a previous experience, whereas know responses refer to a sense of familiarity without contextual detail (Tulving, 1985b). Comparing the neural activity associated with remember and know responses is thus expected to provide additional insight into the substrates of specific types of conscious experiences considered under the general rubric of explicit memory.

In an event-related fMRI study of remembering and knowing (Henson, Rugg, Shallice, Josephs, & Dolan, 1999), subjects first studied a list of words. For each item on a subsequent recognition test, participants responded “remember” or “know” to items they judged to be “old” and otherwise responded “new.” Both correct remember responses and correct know responses, relative to new-correct rejections, were associated with activity in prefrontal cortex (dorsolateral and medial) and medial parietal cortex (precuneus). Relative to new-correct rejections, remember judgments (but not know judgments) were also associated with additional activity in parietal cortex (superior parietal lobule and inferior parietal lobule) and the MTL (parahippocampal gyrus). The direct contrast between remember and know responses complemented these results by showing activity in the parietal cortex (superior parietal lobule and inferior parietal lobule). Although the MTL activation did not survive this direct contrast, it should be noted that only remember responses (versus new-correct rejections) evoked activity in the MTL, providing some indication that this region is preferentially associated with remembering. The reverse contrast between know and remember was associated with activity in the prefrontal cortex (dorsolateral), the parietal cortex (inferior parietal lobule), the MTL (both hippocampus and parahippocampal gyrus), and the fusiform gyrus. This fusiform gyrus activity (coupled with the MTL activity) likely reflects a greater degree of sensory reactivation associated with remember as compared to know responses. The know greater than remember contrast was associated with a distinct region in the (anterior) prefrontal cortex.

In an event-related fMRI remember-know paradigm conducted by Wheeler &
Buckner (2004; adapted from a paradigm originally designed to investigate memory-related domain specific sensory reactivation; see Wheeler & Buckner, 2003; Wheeler et al., 2000), words were paired with either sounds or pictures at study. On the subsequent recognition test, old words (those previously paired with pictures, the only type of old items considered in the analysis) and new words were presented. Participants responded “remember,” “know,” or “new.” Correct remember and know responses were associated with the same degree of activity in one subregion of the parietal cortex, whereas another subregion was associated with greater activity for remember than know responses (both regions were in the inferior parietal lobule). The contrast of remember versus know was also associated with activity in the prefrontal cortex (medial), the MTL (hippocampus), and the fusiform cortex. Because subjects were remembering previously studied pictures, the fusiform cortex activity in this study likely reflects memory-related sensory reactivation. Know versus remember responses were associated with activity in the (dorsolateral) prefrontal cortex.

Across the studies reviewed, a number of patterns emerge (Table 28.1, middle column). First, remembering and knowing, as compared to new-correct rejections, were associated with activity in prefrontal cortex and parietal cortex. Second, remembering evoked greater activity than knowing most consistently in parietal cortex, the MTL, and sensory cortex. Third, knowing evoked greater activity than remembering in the prefrontal cortex. These findings are largely consistent with evidence from remember-know ERP studies, which have shown greater remember than know activity at parietal scalp electrodes (approximately 400–800 ms from stimulus onset) in addition to similar remember and know activity (both greater than new) at frontal scalp electrodes (approximately 1000–1600 ms from stimulus onset; Curran, 2004; Duarte, Ranganath, Winward, Hayward, & Knight, 2004; Düzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Smith, 1993; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999; see also Wilding & Rugg, 1996).

Neuropsychological evidence converges to some extent with the neuroimaging (and ERP) findings. In a study by Knowlton and Squire (1995), amnesic patients with MTL damage studied a list of unrelated words and then made “remember,” “know,” or “new” judgments on a subsequent recognition test. Amnesic patients showed a large decrement in remember responses as compared to control participants and a more modest but still significant decline in know responses (at a 10-minute delay between study and test). Subsequent studies showed a similar pattern of results, where amnesic patients showed a severe impairment in remembering along with more modest trends for impairments in knowing (Schacter, Verfaellie, & Anes, 1997b; Schacter, Verfaellie, & Pradere, 1996c; Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998) and unilateral temporal lobectomy patients have been shown to only be impaired in remembering (Moscovitch & McAndrews, 2002).

Although these group studies include patients with damage to a variety of MTL structures, more recent studies have attempted to distinguish between patients with damage restricted to the hippocampal formation and those with more extensive MTL damage. Yonelinas et al. (2002) found deficits in both remembering and knowing in patients with damage to both the hippocampus and surrounding parahippocampal gyrus. By contrast, they found impairments of remembering – but not knowing – in patients who developed memory deficits as a result of hypoxia, which is known to produce damage restricted to the hippocampal formation in patients whose deficits are restricted to memory (see Yonelinas et al., 2002). Note, however, that anatomical information was not provided concerning the precise lesion sites of the hypoxic patients included in the Yonelinas et al. (2002) study, so the anatomical implications of these findings are uncertain. Manns, Hopkins, Reed, Kitchener, and Squire (2003) reported significant and comparable deficits of remembering and
knowing in amnesics with restricted hippocampal damage, compared with controls. By contrast, a recent case study of patient B.E., who has selective bilateral hippocampal damage, suggests that damage to the hippocampal region alone can result in a specific deficit in remembering with relative sparing of knowing (Holdstock, Mayes, Gong, Roberts, & Kapur, 2005; see also Holdstock et al., 2002).

In summary, although all neuropsychological studies of amnesic patients with MTL or restricted hippocampal damage reveal severe deficits of remembering, the evidence is mixed concerning the role of the MTL generally, and of hippocampus specifically, in knowing. Given current controversies in the interpretation of remember/know data (cf., Dunn, 2004; Gardiner et al., 2002; Rotello, Macmillan, & Reeder, 2004; Wixted & Stretch, 2004), it is perhaps not entirely surprising that clarification of the relative status of remembering and knowing in amnesic patients will require further study. Nonetheless, these neuropsychological studies complement imaging data by providing evidence that the MTL is critically involved in remembering, which reflects a rich form of conscious recollective experience (see discussion in Moscovitch, 1995, 2000).

Sensory Reactivation in True and False Memory

In the type of recognition memory paradigms we have considered thus far, analyses of cognitive and brain activity typically focus on accurate responses: “old” responses to studied items (old-hits) or “new” responses to non-studied items (new correct rejections). False alarms to new items in such paradigms are usually too few to allow meaningful analysis. However, cognitive psychologists have developed a number of paradigms that yield much larger numbers of false alarms, thus allowing comparison of the cognitive and neural properties of true memories and false memories (for a recent review, see Schacter & Slotnick, 2004). For example, Roediger and McDermott (1995), extending earlier work by Deese (1959), reported a paradigm that produces extremely high levels of false memories (now commonly referred to as the DRM paradigm). In the DRM paradigm, participants are presented with lists of associated words (e.g., fly, bug, insect, web, and other related words) that are related to a non-studied lure item (e.g., spider). Roediger and McDermott’s (1995) study showed that subjects falsely recognized a high proportion of these related lure items and often claimed to specifically “remember” (versus “know”) that the lure items appeared on the study list. A similar paradigm has been used to study false memory for visual shapes, in which subjects study physically related shapes and later produce high levels of false alarms to perceptually similar shapes that had not been previously seen (Koutstaal, Schacter, Verfaellie, Brenner, & Jackson, 1999). From the perspective of the present chapter, the development of such paradigms allows us to examine the similarities and differences in the neural correlates of conscious experiences associated with accurate and inaccurate memories.

In the first neuroimaging study to compare true and false memory (Schacter, Reiman, Curran, Yun, Bandy, McDermott, & Roediger, 1996b), participants heard DRM-associated lists followed by a recognition test (consisting of studied/old words, lures/non-studied related words, and non-studied unrelated/new words). Each item type at test (old, related, and new), in addition to a baseline passive fixation condition, was presented in a separate PET scanning block. Both true and false recognition, compared to baseline fixation, were associated with activity that included anterior/dorsolateral prefrontal cortex (BA 10/46), precuneus (medial parietal cortex), and parahippocampal gyrus (within the MTL). The direct contrast between true and false recognition was associated with activity in a left temporal parietal cortex, a region linked to auditory processing. This latter finding can be taken as evidence for greater sensory reactivation (i.e., auditory cortex activation dur-
ing memory for previously spoken words) during true memory as compared to false memory.

A similar experiment was conducted using event-related fMRI (Schacter et al. 1997a), where event types during the recognition test were intermixed, and it showed that both true and false recognition (compared to baseline fixation) were associated with similar patterns of activity including the prefrontal cortex, parietal cortex, and the MTL. However, unlike the previous study, the true greater than false recognition contrast did not reveal activity in any region. At the same time, an ERP experiment suggested that true greater than false recognition-related activity could be attributed to differences in blocked versus event-related designs (Johnson, Nolde, Mather, Kounios, Schacter, & Curran, 1997). Although this latter finding suggests common neural substrates underlying true and false recognition, subsequent fMRI studies have shown more convincing evidence of true/false differences in brain activity and have begun to elucidate the nature of that activity.

In an event-related fMRI study conducted by Cabeza, Rao, Wagner, Mayer, and Schacter (2001), a male or female (on videotape, a relatively rich contextual environment) spoke words from DRM lists of semantic associates or similar categorized lists (e.g., onion, cucumber, and pea are exemplars of the category ‘vegetable’). Participants were instructed to remember each word and whether it was spoken by the male or female. At test, old words, related words (e.g., non-presented associates and categories), or new words were presented, and participants made an old-new recognition decision. True recognition and false recognition, as compared to new items, were associated with activity in the dorsolateral prefrontal cortex, parietal cortex (medial and inferior parietal lobule), and the MTL, specifically the hippocampus. The contrast of false recognition versus true recognition was associated with greater activity in the parietal cortex (inferior parietal lobule) and another region of the MTL, the parahippocampal gyrus. This parahippocampal gyrus activity (which has also been reported in a true/false recognition paradigm by Okado and Stark, 2003) may reflect greater true than false recognition-related contextual reactivation (possibly reflecting memory for the videotaped speakers), because this region has been associated with processing visual context (Bar & Aminoff, 2003; Epstein & Kanwisher, 1998).

Slotnick and Schacter (2004) used event-related fMRI to investigate the neural substrates of true and false recognition for abstract visual shapes. During the study phase, participants viewed sets of exemplar shapes that were similar to a non-presented prototype shape (analogous to DRM word lists). At test, old shapes, related shapes (e.g., non-studied but similar shapes), or new shapes were presented, and participants made an old-new recognition judgment. True recognition and false recognition, as compared to new-correct rejections (i.e., responding “new” to unrelated new items), were associated with activity in the anterior/dorsolateral and medial prefrontal cortex, parietal cortex (superior parietal lobule, inferior parietal lobule, and precuneus), the MTL (hippocampus), and ventral occipital-temporal visual processing regions (BA 17/18/19/37). Although the true greater than false recognition contrast and the reverse contrast were each associated with activity in different regions of the dorsolateral prefrontal cortex and parietal cortex (including precuneus and inferior parietal lobule), only the true greater than false recognition contrast was associated with activity in visual processing regions, specifically in BA 17 and BA 18. These latter regions may reflect greater visual sensory reactivation associated with true recognition as compared to false recognition. The results of ERP studies investigating the neural basis of true and false visual spatial memory are consistent with these findings (Fabiani, Stadler, & Wessels, 2000; Gratton, Corballis, & Jain, 1997).
Thus, across a number of studies, cortical activity that is likely associated with sensory/contextual processing is greater for true than false recognition. Such differential activity might be taken as reflecting conscious recollection of sensory/contextual details that are remembered during true but not false recognition, an idea that has received some support from behavioral studies (e.g., Mather, Henkel, & Johnson, 1997; Norman & Schacter, 1997). Slotnick and Schacter (2004) attempted to identify visual processing regions that reflect conscious memory by contrasting old-hits (responding "old" to old items) and old-misses (responding "new" to old items; see also, Wheeler & Buckner, 2003, 2004). If activity within such regions reflects conscious memory, then brain activity should be greater for old-hits than for old-misses. Conversely, regions that reflect non-conscious memory should respond equivalently during old-hits and old-misses, but in both cases to a greater degree than during new correct rejections (Rugg, Mark, Walla, Sch洛serscheidt, Birch, & Allan, 1998). Slotnick and Schacter found that conscious memory, as identified by the old-hits greater than old-misses contrast, was associated with activity in later visual processing regions (BA 19/37), whereas non-conscious memory — identified by contrasting both old-hits and old-misses each with new-correct rejections — was associated with activity in earlier visual processing regions (BA 17/18). The same functional-anatomic dichotomy was also observed in a follow-up experiment. Both the true greater than false recognition activity and the old-hits and old-misses greater than new-correct rejections results provide convergent evidence that activity in BA17/18 reflects nonconscious memory, at least in the paradigm used by Slotnick and Schacter.

We have proposed that this early visual area activity may reflect the influence of priming, which as noted earlier is a non-conconscious form of memory (Slotnick & Schacter, 2004). One possible problem with this idea is that neuroimaging studies have often shown that priming is associated with decreases in activity following repetition of familiar items, such as words and pictures of common objects (for reviews, see Henson, 2003; Schacter & Buckner, 1998; Schacter et al., 2004; Wiggs & Martin, 1998). However, it has also been found that repetition of novel (or masked) faces, objects, or shapes elicits increases in regional brain activity (e.g., Henson, Shallice, & Dolan, 2000; James, Humphrey, Gati, Menon, & Goodale, 2000; Schacter, Reiman, Uecker, Polster, Yun, & Cooper, 1995; Uecker, Reiman, Schacter, Polster, Cooper, Yun, & Chen, 1997). Because novel abstract shapes served as materials in the Slotnick and Schacter study, the priming hypothesis remains viable. Note also that Slotnick and Schacter used the identical shapes at study and test, thus allowing for repetition priming to occur (see also, Slotnick et al., 2003), whereas other visual memory studies that failed to observe memory-related activity in early visual regions (but found memory-related activity in late visual regions BA 19/37) did not use the identical stimuli at study and test, thus reducing the possibility of repetition priming (Vaidya et al., 2002; Wheeler & Buckner, 2003; Wheeler et al., 2000).

The overall pattern of results thus suggests that memory-related activity in BA 17/18 may be non-conscious. This observation has ramifications for interpreting activity associated with performance on explicit memory tests. Typically, activation associated with explicit memory tests such as old-new recognition is attributed to conscious processing; however, the present analysis indicates this is not always the case. Rather, additional analyses (such as the old-hits versus old-misses contrast) appear necessary to investigate and characterize the nature of activity associated with explicit memory.

The idea that activity in early visual regions that distinguishes between true and false recognition reflects non-conscious memory processes may also help explain why false recognition occurs at high levels, even though brain activity can distinguish between true and false memories. If the activity in early visual regions that distinguished between true and false memories
had been consciously accessible, participants should have used this activity to avoid making false alarms to the related shapes. The fact that there was nonetheless a high rate of false recognition makes sense if the activity within these regions reflects a non-conscious form of memory.

Across the true and false recognition studies reviewed, a number of patterns can be observed (Table 28.1, right column). First, consistent with regions previously associated with explicit retrieval, true and false recognition (versus new-correct rejections) were both associated with activity in prefrontal cortex, parietal cortex, and the MTL (most consistently within the hippocampus). Second, true greater than false recognition was associated with activity in the parietal cortex and sensory/contextual processing regions. Third, false greater than true recognition was associated with activity in the prefrontal cortex (distinct from the commonly active regions).

Neuropsychological studies have provided convergent evidence, particularly regarding the role of the MTL in both true and false recognition. In a study by Schacter et al. (1996c), amnesic patients (with MTL damage) took part in a recognition memory paradigm that used associative word lists. As expected, these patients showed lower levels of true recognition (and higher levels of false alarms to new words) as compared to control participants; in addition, the patients had lower levels of false recognition (i.e., a reduced rate of false alarms to semantically related words; see also, Melo, Winocur, & Moscovitch, 1999; Schacter, Verfaellie, Anes, & Racine, 1998). Similarly, reduced levels of both true and false recognition in amnesic patients have also been shown in recognition memory paradigms that have employed conceptually related words (e.g., “twister,” “funnel”) and perceptually related words (e.g., “hate,” “mate”; Schacter et al., 1997b), or abstract visual patterns (Koutstaal et al., 1999; similar to those used by Slotnick & Schacter, 2004). Furthermore, Alzheimer’s disease patients (with neuropathology that includes, but is not limited to, the MTL regions) also have lower levels of false recognition as compared to control participants (Balota, Watson, Duchek, & Ferraro, 1999; Budson, Daffner, Desikan, & Schacter, 2000; Budson, Desikan, Daffner, & Schacter, 2001; Budson, Sullivan, Daffner, & Schacter, 2003). These neuropsychological studies indicate that the MTL is critically involved in both true and false recognition.

Concluding Comments

In this chapter we have reviewed cognitive neuroscience evidence concerning three distinctions that illuminate different aspects of the relation between memory and consciousness: retrieval success versus attempt, remembering versus knowing, and true versus false recognition. Retrieval success involves memory of a previously experienced item or event, whereas retrieval attempt refers to the effort associated with remembering (without success). As such, successful retrieval (based on the associated memorial experience/details) can be said to reflect high retrieval content. whereas retrieval attempt can be said to reflect low retrieval content. By definition, remember-know studies are used to study distinctions between contextual differences in explicit memory: Remember responses are associated with greater sensory/contextual detail (i.e., high retrieval content), whereas know responses are not associated with sensory/contextual detail (i.e., low retrieval content). True recognition has been associated with access to greater sensory/contextual detail as compared to false recognition (Mather et al., 1997; Norman & Schacter, 1997; Schooler, Gerhard, & Loftus, 1986). Accordingly, retrieval content can be considered greater during true as compared to false memory (although not to such a degree as to preclude the occurrence of false memories). Thus, although both true and false recognition are forms of explicit memory, where common neural substrates likely reflect mechanisms of general retrieval, regions differentially associated with true and false recognition can be assumed to reflect high and low retrieval content, respectively.
As reflected in our summaries at the conclusion of each section of the chapter, the patterns of results for retrieval success and attempt, for remembering and knowing, and for true and false recognition show striking parallels (Table 28.1). The patterns of results for retrieval success and attempt differed from the patterns for remembering and knowing only in that remembering greater than knowing (and not retrieval success greater than attempt) was associated with activity in parietal and sensory cortex (which may simply reflect general differences in the use of stimulus materials; e.g., pictures versus words). The patterns of results for true versus false recognition were largely identical to the patterns of results for remembering versus knowing, except that true and false recognition were both associated with MTL activity, whereas some data indicate remembering but not knowing were associated with MTL activity (as noted earlier, however, the neuropsychological evidence for this conclusion is uncertain, with some data indicating a link between knowing and MTL structures). That the MTL is associated with false recognition may provide some explanation why participants respond “old” despite the fact there may be less contextual detail associated with these items.

We now consider the common neural activity associated with high retrieval content (i.e., retrieval success, remembering, and true recognition) and low retrieval content (i.e., retrieval attempt, knowing, and false recognition). Memories with both high and low retrieval content were associated with activity in the prefrontal cortex and parietal cortex, which indicates these regions are generally associated with explicit retrieval. There was also some evidence that memories with low retrieval content, to a greater degree than those with high retrieval content, may be associated with increased prefrontal cortex activity; however, this activity has been attributed to greater low retrieval content-related post-retrieval monitoring (Schacter & Slotnick, 2004). That is, although there may be more effortful conscious processing with low retrieval content items (which can be considered access-consciousness; see Block, 1995), and is perhaps attributable to greater task difficulty, this is typically not the central focus in discussions of consciousness and memory. Rather, high retrieval content and low retrieval content refer to the sensory/contextual experience associated with retrieval of episodic memories. Relevant to this point, high retrieval content memories, to a greater degree than low retrieval content memories, were associated with activity in the parietal cortex (most consistently the inferior parietal lobule) and sensory processing regions (at least for remembering and true recognition, with a null result for retrieval success). The parietal activity may reflect a greater degree of attention during retrieval of memories with high retrieval content as compared to those with low retrieval content (Corbetta & Shulman, 2002; Hopfinger, Buonocore, & Mangun, 2000). Critically, however, the greater degree of sensory activity associated with memories with high versus low retrieval content provides evidence that memories are constructed by reactivation of features that comprised a previous item or event (Squire, 1992; Schacter et al., 1998a).

The present chapter shows that a cognitive neuroscience approach can illuminate the relation between memory and consciousness, highlighting how explicit memories with different degrees of retrieval content can be linked to distinct neural substrates. Although we would be remiss not to point out that this area of research is in its infancy, we also believe that the field has advanced significantly since the publication of Tulving’s (1985b) lament concerning the lack of interest in memory and consciousness. We suspect that advances during the next 20 years will be even more impressive than those of the past two decades.

Acknowledgments

Preparation of this chapter was supported by grants NIA AG08441 and NIMH MH060951.
THE COGNITIVE NEUROSCIENCE OF MEMORY AND CONSCIOUSNESS

References


Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and...


Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats,


