The Cognitive Neuroscience of Implicit and False Memories:

Perspectives on Processing Specificity

Daniel L. Schacter[1], David A. Gallo[2], and Elizabeth A. Kensinger[1]

[1]Department of Psychology, Harvard University and Athinoula A Martinos Center for Biomedical Imaging

[2]Department of Psychology, University of Chicago

Correspondence to: Daniel L. Schacter, Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138   e-maiL dls@wjh.harvard.edu

The Cognitive Neuroscience of Implicit and False Memories:

Perspectives on Processing Specificity

Roddy Roediger's research career has touched on many aspects of memory, but a careful analysis suggests that it divides into three relatively distinct stages. Stage one, lasting for roughly a decade from 1975-1985, focused on phenomena illustrating that memory performance can be surprisingly good (hypermnesia; e.g., Roediger & Payne, 1982) or perplexingly poor (inhibitory effects of cueing and the act of recall on subsequent memory; e.g., Roediger, 1978). Stage two stretched across the next decade and focused on priming and related implicit memory effects that reflect retention of previously studied information in the absence of conscious recollection (e.g., Roediger, Weldon, & Challis, 1989; Roediger, 1990). The third stage runs from 1995 to the present, and is characterized by a focus on false memories, that is, remembering events that never happened (e.g., Roediger & McDermott, 1995). All three stages have been highly productive, although the fruits of the latter two stages are perhaps better known than those of the first.

The historical record will show that the aforementioned focus on implicit and false memories during stages two and three of Roediger's career parallel to a large extent those that constituted the focus of Schacter's lab at the same time. While Roediger's approach to implicit and false memories has relied largely on cognitive techniques, attempts by Schacter and colleagues to study these phenomena have often adopted a cognitive neuroscience orientation that focuses on the use of neuropsychological and neuroimaging approaches. These latter approaches have been increasingly prominent in Roediger's recent work. In this chapter we pay tribute to Roddy Roediger's career by taking stock of some of the things we have learned about

implicit and false memories as a result of adopting a cognitive neuroscience approach. We focus especially on issues related to the *specificity* of such memories and the processes that support them. It seems especially appropriate to stress the importance of processing specificity when considering both implicit and false memories because this general theme has been important in Roediger and colleagues' ideas concerning transfer appropriate processing (e.g., Roediger, Challis, & Weldon, 1989) and more recently, components of processing (Roediger, Buckner & McDermott, 1999).

The chapter will begin by considering a particular type of implicit memory – priming– that has been the subject of extensive experimental and theoretical debate over the past two decades (e.g., Roediger, 1990; Tulving & Schacter, 1990). We consider some of the modern historical origins of priming research, and then consider recent neuroimaging studies. We will emphasize the specificity of priming – that is, when priming reflects retention of specific features of previously perceived items or specific responses made to those items (Schacter, Dobbins, & Schnyer, 2004). Roediger and colleagues were among the first cognitive psychologists to emphasize the empirical and theoretical importance of specificity for the analysis of priming (e.g., Roediger & Blaxton, 1987), and the issue is turning out to be as central to the concerns of cognitive neuroscience as it has been for cognitive psychology.

We will then turn to the analysis of false memories. Roediger and McDermott's (1995) demonstration of high levels of false recognition of semantic associates has served to focus attention on both the cognitive and neural mechanisms that underlie such robust false recognition effects (for reviews, see Roediger & McDermott, 2000; Schacter & Slotnick, 2004). We will consider research that examines the nature of encoding processes that result in the creation of false memories, with an emphasis on the specificity of processing that supports false recognition.

Equally important, attempts to understand robust false recognition have spawned a line of research exploring the conditions under which people can reduce or suppress such potent memory distortions (e.g., Gallo, 2004; Schacter, Israel, & Racine, 1999). We will consider two pertinent lines of work from our laboratory. One has examined the role of recollecting distinctive information in reducing false memories. We will summarize studies indicating that false memory reduction is mediated by specific and dissociable cognitive mechanisms, rather than by a generalized enhancement of memory. We then turn to research that focuses on the role of emotion in modulating memory distortion. We review cognitive studies indicating that emotional arousal can result in reductions of memory errors. Furthermore, neuroimaging studies indicate that such reductions depend on emotion-specific processes rather than on more generalized memory mechanisms. Thus, the theme of processing specificity will be central to our discussion of false memory reduction.

<p style="text-align:center">Priming: Lessons from Specificity Effects</p>

Priming refers to a change in the ability to identify, produce, or classify an item as a result of a prior encounter with that item or a related item. Modern research on priming can be traced to at least three different though related developments during the late 1970s and early 1980s. First, researchers interested in the nature of lexical processing and word recognition reported priming effects on lexical decision and word identification tasks, which were thought to provide insight into the nature and structure of lexical representation (e.g., Scarbrough, Cortese, & Scarborough, 1977). Second, studies of amnesic patients with damage to the medial temporal lobe and related structures revealed preserved priming effects on tasks such as word stem completion, that do not require conscious recollection of prior experiences, despite patients' severe impairments of explicit memory (e.g., Graf, Squire, & Mandler, 1984; Schacter, 1985;

Warrington & Weiskrantz, 1974). Third, research with healthy young adults revealed experimental dissociations between priming and explicit memory produced by such manipulations as depth of encoding and retention interval (e.g., Jacoby & Dallas, 1981; Graf & Mandler, 1984; Tulving, Schacter, & Stark, 1982).

Almost from the beginning of research on priming, theoretical and experimental interest focused heavily on the specificity of the observed effects, that is, the extent to which priming reflects retention of detailed information acquired during a specific prior episode, versus activation of an abstract representation in long-term memory. Roediger and colleagues were among the first to emphasize the critical theoretical importance of priming specificity in the context of their ideas regarding transfer-appropriate processing (TAP; e.g., Roediger et al., 1989). In several important experimental papers, they showed that priming effects on such tests as word fragment completion show modality specificity (i.e., priming is reduced significantly by study-to-test changes in sensory modality), and also show specificity within a modality (e.g., changes in case or font of a word between study and test can reduce priming; see Roediger & Blaxton, 1987; Weldon & Roediger, 1987).

Since these early studies, numerous experiments have examined the nature and extent of specificity in priming, with much theoretical debate surrounding the aforementioned issue of whether priming depends on nonspecific, abstract lexical or perceptual entities versus specific episodic or exemplar representations (cf., Bowers, 2000; Tenpenney, 1995). Discussion of this extensive literature is beyond the purview of the present chapter. Instead, we wish to focus on a type of specificity that has recently been discovered in the context of neuroimaging research on priming. Schacter et al. (2004) recently proposed a distinction among three types of specificity: stimulus, associative, and response. *Stimulus specificity* occurs when priming is reduced by

changing physical properties of a stimulus between study and test; *associative specificity* occurs when priming is reduced because associations between target items are changed between study and test; and *response specificity* occurs when priming is reduced because subjects make different responses to the same stimulus item at study and test. We consider here response specificity, which has only recently come to the attention of priming researchers. We focus on recent neuroimaging and neuropsychological data that delineate the phenomenon, and consider the theoretical implications of response specificity with respect to Roediger's ideas on TAP and contemporary accounts of brain activity associated with priming.

Neuroimaging of priming and response specificity

Studies that have demonstrated response specificity in priming have been conducted with a view toward testing ideas that have been advanced to explain priming-related changes in brain activity observed in neuroimaging studies using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI). In such studies, participants are scanned while they carry out a task used to assess priming, such as completing three letter word stems with the first word that comes to mind or making judgments about pictures of familiar objects. During primed scans, participants are given target items (e.g., word stems or objects) that appeared previously during the experiment; during unprimed scans, the target items did not appear previously. Virtually all studies using such procedures report *decreased* activity in several cortical regions during primed scans compared to unprimed scans, most consistently in areas within the frontal lobes and the extrastriate visual cortex (for reviews, see Henson, 2003; Schacter & Buckner, 1998; Wiggs & Martin, 1998).

Wiggs and Martin (1998; see also Grill-Spector & Malach, 2001; Schacter & Buckner, 1998) contended that priming-related decreases in human neuroimaging studies may be related to

a phenomenon established in studies of nonhuman primates known as *repetition suppression* (for review and discussion, see Desimone & Duncan, 1995). For instance, in an early series of studies by Desimone and Miller and their colleagues (see Desimone & Duncan, 1995) animals viewed complex visual objects, such as patterns or faces, while experimenters recorded the activity of cells in the lower or inferior temporal (IT) cortex. Repeated exposure to the same stimulus resulted in reduced responses across a substantial proportion of IT cells. Repetition suppression in non-human primates and priming-related decreases in humans could reflect the operation of common underlying mechanisms, which serve to sharpen or "tune" neural object representations with repetition. According to such tuning accounts, neural "tuning" or sharpening is linked with faster and more efficient responding (Wiggs & Martin, 1998).

Although the neural tuning account emerges from studies concerned with perceptual priming of visual objects, it can be extended to other priming-like phenomena. For example, Raichle et al. (1994) reported that generating verbal associates to cue words yielded increased activation in the cingulate, left prefrontal and posterior temporal cortex compared with simple reading of stimuli. Importantly, activation declined with repetition of the verb generation task, and this reduction correlated with reduced reaction times. Consistent with a neural tuning account, the activation reductions and associated reaction time decreases could indicate that semantic analysis of the materials is sharpened or streamlined with repetition. An alternative possibility, however, is that semantic analysis of the repeated stimuli is largely bypassed in favor of rapid retrieval of previous instances that directly indicate the appropriate response. More generally, it is possible that activation reductions in at least some previous priming studies could reflect such response learning rather than neural tuning.

Dobbins, Schnyer, Verfaellie, & Schacter (2004) attempted to directly contrast tuning and response learning accounts with an object decision priming task that had been used in previous neuroimaging research, and yielded evidence of reductions in priming-related activation in regions of prefrontal and fuiform cortex (e.g., Koutstaal et al., 2001). Dobbins et al. modified the task so that responses either remained the same or changed across repeated trials. In the first scanning phase, pictures of common objects were either shown once or repeated three times, and subjects indicated whether each stimulus was bigger than a shoebox using a "yes" or "no" response. In the next phase, the cue was inverted so that subjects were now required to indicate whether each item was "smaller than a shoebox"; they made this judgment about new items, and a subset of those that had been shown earlier. In the final scanning phase, the cue was restored to "bigger than a shoebox" and subjects were tested on new items and the remaining items from the initial phase.

If priming-related reductions in neural activity that are typically produced by this task represent facilitated size processing, attributable to "tuning" relevant aspects of neural representations, then cue reversal should have little effect priming (though it could disrupt overall task performance by affecting both new and primed items). According to the neural tuning account, the same representations of object size should be accessed whether the question focuses on "bigger" or "smaller" than a shoebox. By contrast, if subjects come to rapidly recover prior responses, and this response learning mechanism bypasses the need to recover size representations, then the cue reversal should disrupt priming-related reductions. When the cue is changed, subjects would have to abandon learned responses and instead re-engage the target objects in a controlled manner in order to recover size information.

The fMRI data supported the latter account. During the first scanning phase, standard priming-related activation reductions were observed in both anterior and posterior regions previously linked with priming: left prefrontal, fusiform and extrastriate regions. When the cue was reversed, however, these reductions were eliminated in the left fusiform cortex and disrupted in prefrontal cortex; there was a parallel effect on behavioral response times. But when the cue was restored to the original format, priming-related reductions returned (again there was a parallel effect on behavioral response times), suggesting that the reductions depended on the ability of subjects to use prior responses during trials. Note, however, that the effect was seen most clearly for items repeated three times before cue reversal, a point to which we return shortly.

Finally, the response learning account was also supported by multiple regression techniques demonstrating that reductions in left prefrontal activity predicted the magnitude of behavioral priming for individual subjects: greater initial reductions in prefrontal activity were associated with greater subsequent disruptions of behavioral time. To the extent that prefrontal reductions signal less reliance on controlled processing, and greater reliance on automatic processing, these data suggest that performance disruptions attributable to response reversal reflect a need to re-engage slower controlled processes in order to make object decisions. This idea is consistent with the finding that reductions in fusiform activity do not predict behavioral costs of switching cues, suggesting that these reductions may be incidental to the behavioral facilitation.

This evidence for response specificity in neural priming, as indexed by fMRI signal changes, and in behavioral priming, as indexed by changes in response latencies, is perhaps surprising because previous priming research has neither documented such effects nor even

considered their possible existence (for discussion of prior studies related to response specificity, see Dobbins et al., 2004; Schacter et al., 2004). Clearly, there must be limitations on response specificity: a number of well-established priming effects occur when participants make different responses during study and test. For instance, priming effects on the stem completion task, where subjects respond with the first word that comes to mind in response to a three-letter word beginning, are typically observed after semantic or perceptual encoding tasks that require a different response.

Nonetheless, the existence of response specificity poses a challenge not only for perceptual tuning accounts of priming-related decreases in fMRI signal, but also for theories that explain behavioral priming effects on object decision and related tasks in terms of changes in perceptual representation systems that are thought to underlie object representation (e.g., Schacter, 1990, 1994; Tulving & Schacter, 1990). Such theories make no provisions for response specificity effects. However, the TAP account of priming advanced by Roediger and colleagues (e.g., Roediger et al., 1989, 1998) can easily accommodate such effects. According to the TAP perspective, priming effects are maximized when the same processing operations are performed at study and test. Although this view has emphasized the role of overlapping perceptual operations at study and test to explain priming effects on tasks such as object decision, to the extent that the subject's decision or response is an integral part of encoding operations, then it makes sense that reinstating such operations at test would maximize priming effects.

The foregoing considerations suggest a nice fit between response specificity effects and Roediger's TAP perspective, but there is one further feature of the experimental paradigm that Dobbins et al. (2004) used to produce response specificity that points to a potential complication. Whereas priming in traditional cognitive studies is based on a single study exposure to a target

item, fMRI studies of priming have typically used several study exposures in order to maximize the strength of the fMRI signal. This procedure was followed by Dobbins et al. (2004), who observed that response specificity effects were most robust for items presented three times during the initial phase of the experiment (high primed items), compared with items presented just once (low primed items).

A more recent neuropsychological investigation of response specificity in amnesic patients highlights the potential theoretical importance of results observed for items presented multiple times during the study phase versus only once (Schnyer, Dobbins, Nicholls, Schacter, & Verfaellie, in press). Numerous previous studies have documented that amnesic patients with damage to the medial temporal lobe can exhibit intact priming effects under a variety of conditions, although it is less clear whether amnesics consistently exhibit normal stimulus and associative specificity effects (see Schacter et al., 2004, for discussion). The question of whether amnesic patients exhibit normal response specificity is pertinent to the question of whether the response specificity effect is based on some type of explicit, conscious memory (in which case the effect should be impaired or absent in amnesic patients) or whether it reflects implicit, nonconsicous memory (in which case amnesic patients would be more likely to exhibit the effect). Schnyer et al. (in press) compared amnesics and controls on a variant of the object decision task used by Dobbins et al. (2004). Objects were presented either once (low primed) or thrice (high primed), and then responses either remained the same ("bigger than a shoebox?") or were switched ("smaller than a shoebox?").

Consistent with the trends from the Dobbins et al. (2004) fMRI study noted above, in healthy controls there was evidence of greater response specificity for high primed objects compared with low primed objects. Amnesic patients showed no evidence of response specificity,

demonstrating normal priming (i.e., decrease in response latencies) for low primed items, and impaired priming for high-primed items. That is, healthy controls showed greater priming for high- than low-primed objects in the same response condition, but amnesics failed to show this additional decrease in response latencies.

These results raise the possibility that different mechanisms may be involved in priming for objects presented once versus those presented multiple times. "One-shot" priming on the object decision task may depend primarily on perceptual systems that operate independently of the medial temporal lobe, and thus are preserved in amnesic patients. In neuroimaging experiments, such effects may indeed reflect some type of "tuning" of perceptual systems, independent of the specific responses that subjects make when processing the object. For items presented several times, however, subjects may learn to associate the object with a particular response. Based on the data from amnesic patients, such response learning appears to require the medial temporal lobe.

Although much further work remains to be done to elucidate the properties and mechanisms underlying response specificity, initial data suggesting that response specificity occurs reliably only after several encounters with an object is perhaps more complex than can be readily explained by a TAP account. More generally, priming researchers have previously not paid much attention to possible differences among mechanisms underlying priming based on single versus multiple exposures to an object, but the combined results from Dobbins et al. (2004) and Schnyer et al. (in press) highlight the need to consider and explore these differences.

<u>Creating and Reducing False Memories</u>

In addition to the processing specificity that characterizes priming, a growing body of evidence indicates that highly specific processes contribute to the expression of false memories. Here we focus on false recognition, or the acceptance of nonstudied stimuli (lures) on a memory

test that is attributable to familiarity or source confusions. In their seminal paper, Roediger and McDermott (1995) found that studying lists of associates (e.g., bed, rest, awake…) led to frequent false recall and recognition of nonstudied associates (e.g., sleep), often with high confidence and an illusory sense of recollection (i.e., the Deese-Roediger-McDermott, or DRM, task). In order to explain these types of false memories, Roediger and colleagues have developed an activation/monitoring framework (e.g., Gallo & Roediger, 2002; McDermott & Watson, 2001; Roediger, Balota, & Watson, 2001; Roediger, Watson, McDermott, & Gallo, 2001). Generally speaking, "activation" refers to those processes that make nonstudied items familiar as a result of associative or conceptual relationships with studied items, whereas "monitoring" refers to those processes that operate on retrieved information to regulate the accuracy of the memory decision.   In the next two sections we review evidence for specificity in both activation and monitoring processes.

Activation specificity

As previously discussed, Roediger and colleagues have developed the TAP framework as one way to understand implicit memory phenomena, with the amount of priming in a task depending on the match between encoding and retrieval processes. Priming studies also indicate that these principles apply to the activation processes that can cause false memories. Several studies have found that presenting DRM lists leads to priming of nonstudied associates on a variety of implicit memory tests (e.g., fragment or stem completion, McDermott, 1997; McKone & Murphy, 2000; Smith, Gerkens, Pierce, & Choi, 2002; lexical decision, Whittlesea, 2002; Hancock, Hicks, Marsh, & Ritschel, 2003; anagram solution, Lövdén & Johansson, 2003). These priming effects indicate that an abstract or lexical representation of related lures can be activated during the presentation of the study list (e.g., McDermott, 1997) and/or that semantic

gist is activated or encoded (e.g., Verfaellie, Page, Orlando & Schacter, in press), giving rise to subsequent priming effects on a variety of tests that are sensitive to the activation of these types of conceptual information.

According to TAP, one would expect that priming for studied words would be more sensitive to perceptual manipulations than would priming for nonstudied associates, because only the studied words were actually perceived during encoding. Hicks and Starns (2005) have recently provided evidence that supports this hypothesis. In their first experiment, they found that priming on a visual word-stem completion test was greater for list words that were studied visually (.14) than auditorily (.08), demonstrating that priming was enhanced by perceptual overlap between study and test. Significant priming also was obtained for nonstudied associates, but importantly, these effects were equivalent in the visual (.06) and auditory (.07) tests. This pattern suggests that nonstudied associates were primed by the activation of abstract or conceptual information, but that perceptually-specific information was not encoded for these items. In a second experiment, a test that was thought to be less sensitive to conceptual information and more sensitive to perceptual information was used (perceptual identification). Unlike the stem-completion results, only studied words were primed on perceptual identification, again suggesting that memories for nonstudied associates did not contain perceptually detailed information.

The idea that true memories contain more perceptual detail than false memories, but similar levels of conceptual or associative information, is consistent with studies that have measured the subjective content of true and false memories (e.g., Gallo & Roediger, 2003; Norman & Schacter, 1997; Mather, Henkel & Johnson, 1997), and also with neuroimaging studies of the DRM task. Using PET and auditory study conditions, Schacter, Reiman, et al.

(1996) found that true recognition was more likely than false recognition to elicit activity in auditory-specific regions (e.g., superior temporal gyrus and supramarginal gyrus), using fMRI and visual study conditions, Cabeza, Rao, Wagner, Mayer, and Schacter (2001) found that true recognition was more likely than false recognition to elicit activity in areas thought to be related to visual processing (e.g., left parahippocampal gyrus), and using ERP, Fabiani, Stadler, and Wessels (2000) found that true but not false recognition elicited laterality effects that corresponded to the lateralized visual presentation that was used at encoding. Collectively, these results indicate that perceptually-related activity at retrieval is greater for true than false recognition, potentially reflecting the recollection or priming of perceptual details (see Slotnick & Schacter, 2004), and that the nature of this retrieval activity for true recognition is specific to the type of perceptual processes engaged during study (e.g., Wheeler, Petersen, & Buckner, 2000). By comparison, true and false recognition have been found to elicit similar activity in medial temporal regions, including the hippocampus (e.g., Cabeza et al., 2001; Schacter et al., 1996), perhaps reflecting the retrieval of common conceptual or associative information that can drive both true and false recognition.

The TAP framework also is relevant to explicit memory performance in the DRM task. Indeed, TAP was originally developed by Morris, Bransford, and Franks (1977) as an alternative way of understanding "levels of processing" effects on explicit tests such as recognition (e.g., Craik & Lockhart, 1972; Craik & Tulving, 1975). The basic idea was that, contrary to a strong version of the levels framework, encoding tasks that encourage "deep" or semantic processing do not necessarily lead to better memory performance than "shallow" or surface-level processing. Instead, as is the case in implicit memory tasks, the level of performance was thought to depend on the overlap between encoding and retrieval processes. In support of this view, Morris et al.

(1977) demonstrated that shallow processing (e.g., rhyme encoding task) yielded greater accuracy than deep processing (e.g., a semantic encoding task) on a test that used surface-level memory cues (e.g., rhyme recognition). By this view, the typical levels-of-processing effect (deep > shallow) arises simply because semantics happens to be the ideal dimension upon which words can be differentiated (see Roediger & Gallo, 2001, for additional discussion).

Evidence for this same type of processing specificity can be found in the false memory literature, too. Early studies using the DRM task found that, much like true memory, false memory was greater following deep (e.g., semantic) encoding tasks relative to shallow (e.g., phonological) tasks (Thapar & McDermott, 2001; Toglia, Neuschatz, & Goodwin, 1999; Rhodes & Anastasi, 2000). These findings suggest that semantic processing enhanced activation of the related lure, by facilitating the processing of common semantic associations and/or conceptual gist. But what would happen if a levels-of-processing manipulation were used on lists of phonological associates, instead of the more typical lists of semantic associates used in the DRM task (e.g., Schacter, Verfaellie & Anes, 1997; Sommers & Lewis, 1999)? According to a TAP account, focusing on the sound of these types of associates (at encoding) should lead to greater false recognition of phonologically related lures (at test), because the overlapping features between study and test (and the corresponding activation processes) are phonological.

Building on earlier false recognition work (e.g., Coltheart, 1977; Elias & Perfetti, 1973), Chan, McDermott, Watson, and Gallo (in press) crossed two types of lists (semantic or phonological associates) with two types of encoding processes (focusing on meaning or sound). Replicating the typical effect with semantically related lists, meaningful processing led to significantly greater false recognition of semantic associates than did sound processing (means = .64 and .54, respectively). More important, this effect was reversed for phonological lists, with

sound processing leading to significantly greater false recognition of phonologically related lures (means = .45 and .51). Much like the priming specificity effects discussed earlier, this cross-over interaction demonstrates that the activation of related lures is highly specific, with meaningful processing preferentially enhancing activation for semantic associates, and sound processing preferentially enhancing activation for phonological associates. These processing differences also have been linked to differences in brain activity, with the processing of semantically related lists activating a more anterior region of left inferior frontal gyrus (near BA44/45) than the processing of phonologically related lists (near BA6/44; McDermott, Petersen, Watson, & Ojemann, 2003). In contrast to the idea of a processing continuum (from shallow to deep), these component processes appear to be subserved by distinct neurocognitive mechanisms.

In sum, both behavioral and neuroimaging data indicate that the activation processes giving rise to false memories are highly specific. In the DRM task, priming studies, subjective measures, and neuroimaging results all converge on the idea that memories for nonstudied associates contain as much conceptual information, but are less likely to contain perceptual information, than studied associates. Further, within the conceptual domain, the corresponding activation processes are specific to the type of materials (e.g., semantic or phonological). These findings are not surprising from a TAP perspective, which considers all memory performance in terms of the relative overlap between a variety of specific study and test processes, and from the related components-of-processing framework, which stipulates that different processes are potentially subserved by different neurological mechanisms.

<u>Monitoring specificity</u>

In addition to activation processes, there also is evidence that recollection-based monitoring or editing processes that can reduce false memory are highly specific. In this case we are using the term "specificity" in a different sense. When discussing activation processes, and also in the section on priming specificity at the outset of this chapter, we used the term "processing specificity" to refer to the overlap between specific study and test processes, and their neurological underpinnings. The monitoring processes discussed here operate outside the purview of this framework, because they influence memory performance only at the time of retrieval, and so it makes little sense to discuss them in terms of the TAP framework.

In terms of recollection-based monitoring, processing specificity refers to the idea that there are different types of monitoring, depending on different expectations and strategies, and that these processes are highly specific to the materials and task demands. This idea stands in contrast to the more general notion of "discrimination," as might be found in various signal-detection models of recognition memory. According to those frameworks, false recognition results from confusions between target and lure distributions, or the degree of overlap between the two distributions. Manipulations that selectively strengthen the target distribution (or true memory) would enhance discrimination, ideally leading to enhanced true recognition and decreased false recognition (depending on the setting of the response criterion). The classic mirror effect provides empirical support for this type of theory: across a variety of manipulations, increased hits tend to be associated with decreased false alarms (e.g., Glanzer & Adams, 1990).

As discussed by Greene (xxxx, this volume), a reevaluation of the mirror effect suggests that there are enough exceptions to the effect to question its use as a universal principle, and recent false memory data raise additional doubts about the utility of the concept. One reason to

suspect that monitoring processes are more specific (or complicated) than a general discrimination account would suggest is that the relationship between true and false recognition is highly variable in the DRM task. Numerous manipulations have been found to reduce false memory, ostensibly by enhancing recollection-based monitoring, but only some lead to increases in true memory (e.g., repeating study words or slowing presentation rate; Benjamin, 2001; Gallo & Roediger, 2002; McDermott & Waston, 2001; Seamon, Luo, Schwartz et al., 2002). Other manipulations that are thought to enhance monitoring do not necessarily affect true recognition, such as presenting studied words in more distinctive formats (e.g., pictures, Schacter, Israel & Racine, 1999, modality, Smith & Hunt, 1998, and writing or vocalizing studied words, Dodson & Schacter, 2001, Seamon, Goodkind et al., 2003).

Of course, the failure to find a mirror effect in these latter cases might simply have resulted from the use of a more conservative response criterion in the more distinctive conditions, thereby offsetting increases in true recollection. More convincing evidence that different monitoring processes are involved comes from aging studies. Relative to younger adults, older adults are impaired in their ability to use repetition to reduce false recognition (e.g., Benjamin, 2001; Kensinger & Schacter, 1999; Watson, McDermott, & Balota, 2004), but they are just as likely as younger adults to reduce false recognition following distinctive study manipulations (e.g., Dodson & Schacter, 2002; Schacter, Israel & Racine, 1999). Findings like these suggest that qualitatively different types of monitoring processes, or strategic uses of recollection, are involved. For instance, study rate or repetition manipulations would allow more time or opportunities for subjects to determine the critical nonpresented word, and to subsequently reject this item by recalling it as "nonstudied" (a recall-to-reject process). Distinctiveness manipulations, on the other hand, might allow subjects to form more detailed

recollective expectations and therefore avoid falsely recognizing new lures that do not elicit such detailed recollections (a distinctiveness heuristic).

As argued by Gallo (2004) and Gallo, Bell, Beier, and Schacter (in press), these recollection-based monitoring processes are characterized by qualitatively different types of decision process. Whereas recall-to-reject relies on the recollection of information that disqualifies the lure as having occurred (an exclusion strategy), the distinctiveness heuristic is based on the failure to recall expected information, and therefore is only diagnostic of nonoccurrence. For reasons that are not altogether clear, aging may have greater effects on one type of monitoring than on the other. Unfortunately, research in the DRM task has only provided limited clues, because monitoring processes can occur both during the study and test phase of these experiments. For this and other reasons, monitoring processes are difficult to isolate with this task.

To further investigate these issues we have developed the criterial recollection task (e.g., Gallo, Weiss, & Schacter, 2004, Gallo, Kensinger, & Schacter, in press). In brief, subjects studied a list of unrelated red words and pictures, and memory was tested under different sets of instructions (using black words as cues). On the picture test, subjects responded "yes" only if the test word corresponded to a studied picture (rejecting test words that corresponded to a red word), and vice versa on the red word test. Critically, some of the studied items were presented as both red words and pictures, so that the recollection of one format was independent (i.e., did not disqualify) the item as having occurred in the other format. As a result, subjects had to carefully search their memory for recollection of the criterial format.

As shown in Figure 1, Gallo et al. (2004) found that subjects were less likely to make false recognition errors on the picture test than on the red word test, and these effects were

obtained regardless of whether true memory was greater for pictures (due to a picture superiority effect, Experiment 1) or for red words (due to repetition of these stimuli at study, Experiment 2). The fact that false recognition suppression was independent of true memory differences is difficult to reconcile with a unitary strength-based explanation, and instead indicated that subjects had used a recollection-based distinctiveness heuristic. Subjects expected qualitatively different types of recollections from the different formats, and because picture recollections were more distinctive, false recognition was lower on the picture test than on the red word test.

In a more recent aging study, we contrasted these original conditions to a condition where a recall-to-reject process also would be possible. In that condition, studied items were only presented in one format or the other, so that a mutually exclusive rejection strategy was possible at test (i.e., if a picture could be recalled, then the subjects could be sure to reject an item as having been studied as a red word). In younger adults we replicated our original findings (minimal false recognition on the picture test), and similar effects were found in older adults, demonstrating a robust distinctiveness heuristic in both groups. However, false recognition errors on the red word test varied across the two groups. In both groups, exclusion errors were reduced when the recall-to-reject strategy was possible, relative to when it was not, but younger adults were more effective at using this strategy than were older adults. Thus, older adults were just as likely as younger adults to use the absence of picture recollections to monitor memory (i.e., the distinctiveness heuristic), but were impaired in their ability to use the presence of picture recollections to monitor memory (i.e., the recall-to-reject strategy). Much like the DRM findings discussed above, these findings suggest that the two types of recollection-based monitoring processes operate in qualitatively different ways. There also is some evidence to suggest that they are subserved by different neurological mechanisms, but the two have yet to be

directly compared in a single neuroimaging study (see Gallo, Kensinger, & Schacter, in press, for relevant findings and discussion).

In sum, a growing body of evidence suggests that recollection can reduce false recognition in at least two different ways: recall-to-reject or the distinctiveness heuristic. These monitoring processes depend on qualitatively different types of decisions, and they can be experimentally dissociated, arguing against a more general explanation of false recognition reduction in terms of enhanced discrimination. This is not to say, though, that signal detection theories are irrelevant, and indeed multiple-dimensional SDT might provide a useful way to conceptualize the distinctiveness heuristic, as would any theory that allows for qualitatively different types of recollective expectations to influence memory decisions (e.g., Johnson & Raye, 1981). Once these points are acknowledged, the possibility emerges that other types of stimulus distinctiveness, each with highly specific neurocognitive underpinnings, also could reduce false recognition. In the next section we explore this possibility in a very different research domain - emotional distinctiveness.

<center>Emotion-Specific Processing Reduces Memory Distortion</center>

"The attention which we lend to an experience is proportional to its vivid or interesting character; and it is a notorious fact that what interests us most vividly at the time is, other things equal, what we remember best. An impression may be so exciting emotionally as almost to leave a scar upon the cerebral tissues"

– William James, 1890, The Principles of Psychology

Many people share the intuition that when an event contains emotional importance, we remember that event with tremendous detail and indelible accuracy. The term "flashbulb memory," coined by Brown & Kulik (1977), captures the picture-like vividness that individuals often believe to accompany their memories of highly surprising and emotionally evocative events. Despite these intuitions, however, empirical data indicate that memories for emotional events are vulnerable to distortion: Individuals often change their reports about how they first learned of an emotional event while maintaining high confidence in their memories (e.g., Neisser & Harsch, 1992; Weaver, 1993; Schmolck, Buffalo, & Squire, 2000). These studies demonstrate convincingly that emotional memories can undergo distortion. They do not, however, speak to whether emotional content influences the frequency with which memory distortion occurs: Are memories for emotional events less prone to distortion than memories for nonemotional events?

As discussed in the preceding section, the likelihood of falsely endorsing an item as studied can relate to how distinctively an individual encodes, and thus expects to retrieve, the information (reviewed by Schacter & Wiseman, in press). While distinctiveness typically has been manipulated by inclusion of pictorial associations (Dodson & Schacter, 2002; Israel & Schacter, 1997; Weiss et al., 2002; Gallo et al., 2004) or by asking people to generate items (Dodson & Schacter, 2001), emotion may provide another dimension along which an item's distinctiveness can be enhanced. Emotionally arousing experiences include numerous aspects that are not present to the same extent in nonemotional ones (e.g., personal relevance and physiological response; Christianson & Engelberg, 1999; LeDoux, 2002). Consistent with the hypothesis that these additional dimensions increase the distinctiveness of emotional information, a number of studies have demonstrated that individuals are more likely to claim that they vividly remember emotionally arousing items as compared to nonemotional ones (e.g.,

Dewhurst & Parry, 2000; Kensinger & Corkin, 2003; Ochsner, 2000). As studies of flashbulb memories have shown, however, these subjective intuitions about the vividness of a memory do not necessarily map on to objective evaluations of its accuracy. Thus, the critical question is whether this subjectively vivid encoding also corresponds with a reduction in memory distortion.

In a number of studies, the answer seems to be affirmative. For example, after studying a list of words orthographically associated with a nonpresented lure word, individuals were less likely to falsely recall or to falsely recognize the lure word if it was emotional (e.g., to endorse *rape* after studying *cape, nape, tape…*) than if it was neutral (e.g., *hook* after studying *book, look, cook*; Pesta, Murphy, & Sanders, 2001; Kensinger & Corkin, 2004a). The likelihood of falsely recognizing the emotional lures also was related to manipulations in distinctiveness: Including other emotional items at encoding (thereby decreasing the distinctiveness of the emotional lure words) increased the likelihood of false recognition (Pesta, Murphy, & Sanders, 2001). These data suggest that individuals can use the distinctiveness provided by emotion to reduce their rates of false recognition. An important caveat, however, is that in this paradigm the majority of the studied items were neutral. Thus, the emotional lures may have been rejected due to their conceptual incongruence with the studied items, rather than because of specific effects of emotion. These studies, therefore, do not provide strong evidence that emotion, per se, alters the frequency of memory distortion.

More direct evidence for effects of emotional salience on memory distortion has come from examinations of reality monitoring ability (the ability to distinguish what has been perceived from what has been imagined; Johnson & Raye, 1981). Assignment to an external or internal source can typically be made by remembering the types of information associated with an event: Experienced events usually include more sensory, contextual, and semantic

information, whereas imagined events often require more cognitive operations to support the internal generation of information (Johnson & Hirst, 1993; Johnson & Raye, 1981). Remembering these characteristics typically allows accurate memory attributions, although reality-monitoring errors sometimes can occur. Kensinger & Schacter (in press b) demonstrated that these errors occurred less frequently for negative arousing items compared to nonemotional ones. This enhanced discrimination for negative arousing items existed whether or not participants were informed about the memory task that would follow, and the effect was present for both verbal stimuli and single objects.

This finding is consistent with evidence indicating that emotional content enhances the binding of many types of item and event details. For example, emotionally arousing items are more likely to be remembered with details such as the color of font in which a word was written or the location of a word on a screen (D'Argembeau & Van der Linden, 2004; Doerksen & Shimamura, 2000; Kensinger & Corkin, 2003; MacKay et al., 2004). Enhanced binding could combat reality-monitoring errors in a number of ways. It could increase the likelihood that participants remember the critical event details needed to distinguish presented from imagined events (e.g., sensory, perceptual, semantic details; Johnson et al., 1993). It also could prevent stimulus confusion. For example, after studying a picture of an apricot, a person may later falsely believe that they saw a picture of an orange if they remember only general information about that object (e.g., an orange-colored fruit). In contrast, this confusion may result less frequently for emotional items if they tend to be remembered with more detail, allowing them to be distinguished from other items. A recent experiment has provided evidence that individuals are more likely to remember the specific visual details of emotional items compared to nonemotional ones. Kensinger, Garoff-Eaton, & Schacter (in press) presented participants with

colored objects at study, some of which were emotional and some of which were neutral (e.g., a snake, a barometer). At retrieval, they showed participants some objects that were identical to those that had been studied, some that were similar but not identical to studied objects (e.g., a barometer that differed from the one presented at study in features such as color, size, shape, or orientation) and other objects that were novel. They found that emotional content enhanced the likelihood that specific visual details were remembered: Individuals were more likely to indicate correctly that an item was identical to the one studied earlier if it was emotional than if it was neutral.

These behavioral studies demonstrate that negative emotional content can decrease the likelihood of memory distortion. They do not, however, clarify whether this increased memory accuracy for the negative arousing information is specifically related to the processing of the emotional information, or whether it stems from engagement of the same processes that reduce memory distortion for neutral information. As discussed earlier in this chapter, increases in true recollection can decrease false memories via a variety of processes.

Neuroimaging provides one method for examining whether the effects of emotion on memory accuracy result from domain-general processes that serve to boost memory accuracy for emotional and nonemotional information, or from emotion-specific processes. To examine this question, Kensinger & Schacter (in press a) used fMRI to examine the neural processes that were engaged during accurate encoding of emotional and neutral information (i.e., that reduced the likelihood of subsequent reality-monitoring errors). We used a paradigm that had been shown to elicit high levels of reality-monitoring errors (see Gonsalves & Paller, 2000). Participants were scanned as they formed mental images of named objects and indicated whether those objects were bigger or smaller than a shoebox. For half of the items, participants were then shown the

corresponding photo. After a delay, participants performed a recognition task in which they had to indicate whether items had been shown as photos. In some instances, reality-monitoring errors occurred, and participants indicated that a photo of an object had been studied, when the participant had only imagined it. As in the behavioral studies discussed above, reality-monitoring errors occurred less frequently for the emotional items than for the neutral items. The critical question was whether distinct neural processes influenced the frequency of reality-monitoring errors for emotional and neutral items.

The study revealed some overlap in the neural processes that were related to successful encoding for emotional items and neutral items. Most notably, activity in the left posterior hippocampus showed a relation to accurate encoding for both item types (i.e., greater activity during the encoding of items later correctly attributed to a presented or nonpresented source than during the encoding of items later misattributed). Importantly, there also were distinctions in the neural processes based on the emotional content of the items: Enhanced activity in the amygdala and the orbitofrontal cortex corresponded with a reduction in the likelihood of memory misattributions specifically for the emotional items. Activity in these regions showed no relation to memory accuracy for the neutral items. Critically, these are regions that are engaged during the processing of emotional information across a range of paradigms (Bechara et al., 2000; Phan et al., 2002; Zald, 2003). Thus, it appears that the way in which emotional information is processed results in a memory accuracy advantage. Part of the effect of these emotion-specific processes appeared to be exerted via their interactions with regions that promote accurate encoding of both emotional and nonemotional items. Activity in the amygdala was highly correlated with activity in the hippocampus during the encoding of emotional items later accurately attributed (see Dolcos et al., 2004; Kensinger & Corkin, 2004b; Richardson et al.,

2004; Kilpatrick & Cahill, 2003 for other evidence of amygdala-hippocampus interactions).

Thus, the results provide strong evidence that the enhancement in memory accuracy for emotional items did not stem solely from the additional engagement of domain-general processes that enhance accuracy for all items. Rather, domain-specific processes (in the amygdala and orbitofrontal cortex), engaged during encoding, served to enhance memory accuracy for the emotional items.

In a separate experiment, Kensinger & Schacter (in press b) examined whether distinct retrieval processes also influenced memory accuracy for emotional and neutral items. Participants performed the same reality-monitoring task described above, but this time they were scanned during the retrieval phase (i.e., as they indicated whether or not a photo of an object had been studied). The results paralleled those of encoding. Activity in the left hippocampus (this time in an anterior region) was related to accurate retrieval for both emotional and neutral items. Activity in emotion-processing regions (the right amygdala and left orbitofrontal cortex), quite similar to those implicated in accurate encoding, corresponded with accurate retrieval only for the emotional items and not for the neutral items. Interestingly, activity in lateral prefrontal regions and the right posterior hippocampus corresponded with accurate retrieval only for the neutral items. These results demonstrate that the emotional content of items affects the neural processes that are associated with accurate retrieval. Moreover, emotion-specific processing (in the amygdala and orbitofrontal cortex) appears to be linked both to the successful encoding and accurate retrieval of emotional information.

Taken together, these studies demonstrate that information with emotional content is less likely to be remembered in a distorted fashion than information lacking emotional salience. The results help us to understand how emotion can modulate memory performance, highlighting roles

of the amygdala and orbitofrontal cortex at both encoding and retrieval. They also emphasize the fact that false memory reduction does not always result from domain-general processes. Rather, it is clear that when emotional information is more accurately remembered than neutral information, it is because of the engagement of emotion-specific processes. Future research will be needed to examine the range of stimuli for which domain-specific versus domain-general processes relate to reductions in memory distortion.

Concluding Comments

In this chapter we have reviewed evidence for what we term processing specificity in three different domains: priming, false recognition, and emotional memory. Our review of priming research noted the existence of three types of specificity: stimulus, associative, and response. We reviewed recent neural and behavioral evidence documenting response specificity in priming, and considered implications for various theoretical accounts of priming-related activations decreases observed in fMRI studies, as well as cognitive perspective on the processes and systems that support priming. Our discussion of false recognition summarized both behavioral and neuroimaging evidence for a high degree of specificity in the activation processes that give rise to false memories, and we also considered additional evidence for specificity in monitoring processes that can be used to reduce false memories. Finally, in our discussion of emotion and memory distortion we considered neuroimaging evidence that emotion-specific processes are involved in reducing the incidence of reality monitoring errors.

We have grouped a number of phenomena under the rubric of "processing specificity", even though it is not clear how or even whether they are connected. For example, the various forms of priming specificity generally refer to the specificity of various kinds of knowledge or stored representations, our discussion of false memory reduction focused on monitoring

operations that could be applied in principle to different kinds of knowledge, and our discussion

of emotional memory considered specific versus general processes at the level of a domain

(emotional vs. non-emotional). Although the relations among the various types of specificity

considered here remain to be elucidated, it seems clear that questions concerning the level and

degree of processing specificity have come to the fore in many sectors of memory research. We

therefore believe that future empirical and theoretical analyses that explore further the various

forms of processing specificity will likely prove important in enhancing our understanding of

memory.

References

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex, 10*, 295-307.

Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 941-947.

Bowers, J. S. (2000). The modality-specific and -nonspecific components of long-term priming are frequency sensitive. *Memory & Cognition, 28,* 406-14.

Brown, R. & Kulik, J. (1977). Flashbulb memories. *Cognition, 5,* 73-99.

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can medial temporal lobe regions distinguish true from false? An event-related functional MRI study of veridical and illusory recognition memory. *Proceedings of the National Academy of Sciences, 98,* 4805-4810.

Chan, J. C. K., McDermott, K. B., Watson, J. M., and Gallo, D. A. (in press). The importance of material-processing interactions in inducing false memories. *Memory & Cognition.*

Christianson, S-A, & Engelberg, E. (1999). Organization of emotional memories. In T. Dalgleish & M. Power (Eds.), <u>The Handbook of Cognition and Emotion</u> (pp. 211-227). Chichester, Sussex: John Wiley & Sons, Ltd.

Coltheart, V. (1977). Recognition errors after incidental learning as a function of different levels of processing. *Journal of Experimental Psychology: Human Learning & Memory, 3,* 437-444.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11,* 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104,* 268-294.

D'Argembeau, A. & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion, 4,* 173-188.

Desimone, R. and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193-222.

Dewhurst, S. A., & Parry, L. A. (2000). Emotionality, distinctiveness, and recollective experience. *European Journal of Cognitive Psychology, 12,* 541-551.

Dobbins, I. G., Schnyer, D. M., Verfaellie, M, & Schacter, D.L. (2004). Cortical activity reductions during repetition priming can result from rapid response learning. *Nature, 428,* 316-319.

Dodson, C. S. & Schacter, D. L. (2002a). Aging and strategic retrieval processes: Reducing false memories with a distinctiveness heuristic. *Psychology and Aging, 17,* 405-415.

Dodson, C. S. & Schacter, D. L. (2002b). When false recognition meets metacognition: The distinctiveness heuristic. *Journal of Memory and Language, 46,* 782-803.

Dodson, C. S., & Schacter, D. L. (2001). "If I had said it I would have remembered it": reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin and Review, 8,* 155-161.

Doerksen, S., & Shimamura, A. (2001). Source memory enhancement for emotional words. *Emotion, 1,* 5-11.

Dolcos, F., LaBar, K. S, & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron, 5,* 855-63.

Elias, C. S., & Perfetti, C. A. (1973). Encoding task and recognition memory: The importance of semantic encoding. *Journal of Experimental Psychology, 99,* 151-156.

Fabiani, M., Stadler, M. A., & Wessels, P. M. (2000). True but not false memories produce a sensory signature in human lateralized brain potentials. *Journal of Cognitive Neuroscience, 12,* 941-949.

Gallo, D. A. (2004). Using recall to reduce false recognition: Diagnostic and disqualifying monitoring. Journal of Experimental Psychology: *Learning, Memory, & Cognition, 30,* 120-128.

Gallo, D. A., Bell, D. M., Beier, J. S., & Schacter, D. L. (in press). Two types of recollection-based monitoring in young and older adults: Recall-to-reject and the distinctiveness heuristic. *Memory.*

Gallo, D. A., Kensinger, E. A., & Schacter, D. L. (in press). Prefrontal activity and diagnostic monitoring of memory retrieval: fMRI of the criterial recollection task. *Journal of Cognitive Neuroscience.*

Gallo, D. A., & Roediger, H. L., III. (2003). The effects of associations and aging on illusory recollection. *Memory & Cognition, 31,* 1036-1044.

Gallo, D. A., & Roediger, H. L., III. (2002). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory & Language, 47,* 469-497.

Gallo, D. A., Weiss, J. A., & Schacter, D. L. (2004). Reducing false recognition with criterial recollection tests: Distinctiveness heuristic versus criterion shifts. *Journal of Memory & Language, 51,* 473-493.

Glanzer, M., & Adams, J.K. (1990). The mirror effect in recognition memory: data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 5-16.

Gonsalves, B. & Paller, K.A. (2000). Neural events that underlie remembering something that never happened. *Nature Neuroscience, 3,* 1316-1321.

Graf, P., & Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning and Verbal Behavior, 23,* 553-568.

Graf, P., Squire, L. R. & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 164-178.

Grill-Spector, K. & Malach, R. (2001). Fmr-adaption a tool for studing the functional properties of neurons. *Acta Psychologica, 107,* 293-321.

Hancock, T. W., Hicks, J. L., Marsh, R. L., & Ritschel, L. (2003). Measuring the activation level of critical lures in the Deese-Roediger-McDermott paradigm. *American Journal of Psychology, 116,* 1-14.

Henson, R. N. (2003). Neuroimaging studies of priming. *Prog Neurobiol, 70,* 53-81.

Hicks, J. L., & Starns, J. J. (2005). False memories lack perceptual detail: Evidence from implicit word-stem completion and perceptual identification tests. *Journal of Memory & Language, 52,* 309-321.

Israel, L. & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review, 4,* 577-581.

Jacoby, L. L. & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110,* 306-340.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28.

Johnson, M. K. & Hirst, W. (1993). MEM: Memory subsystems as processes. In A.F. Collins, S.E. Gathercole, M.A. Conway, & P.E. Morris (Eds). Theories of Memory (pp. 241-286). East Sussex, England: Erlbaum.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67-85.

Kensinger, E. A., & Corkin, S. (2004a). The effects of emotional content and aging on false memories. *Cognitive, Affective, and Behavioral Neuroscience, 4,* 1-9.

Kensinger, E. A. & Corkin, S. (2004b). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences, USA, 101,* 3310-3315.

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory and Cognition, 31,* 1169-1180.

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D.L. (in press). Memory for specific visual details can be enhanced by negative arousing content. *Journal of Memory and Language.*

Kensinger, E.A. & Schacter, D.L (in press a). Emotional content and reality-monitoring ability: FMRI evidence for the influence of encoding processes. *Neuropsychologia.*

Kensinger, E. A., & Schacter, D. L. (in press b). Retrieving accurate and distorted memories: Neuroimaging evidence for effects of emotion. *NeuroImage.*

Kensinger, E. A., & Schacter, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology, 16*, 399-415.

Kilpatrick, L. & Cahill, L. (2003). Amygdala modulation of parahippocampal and frontal regions during emotionally influenced memory storage. *NeuroImage, 20,* 2091-2099.

Koutstaal, W., Wagner, A. D., Rotte, M., Maril, A., Buckner, R. L., & Schacter, D. L. (2001). Perceptual specificity in visual object priming: fMRI evidence for a laterality difference in fusiform cortex. *Neuropsychologia, 39*,184-199.

LeDoux, J. E. (2002). Cognitive-emotional interactions: Listen to the brain. In R.D. Lane and L. Nadel (Eds). Cognitive Neuroscience of Emot ion. New York, NY: Oxford University Press.

Lövdén, M., & Johansson, M. (2003). Are covert verbal responses mediating false implicit memory? *Psychonomic Bulletin & Review, 10,* 724-729.

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: evidence from taboo stroop, lexical decision, and immediate memory tasks. *Memory and Cognition, 32,* 474-488.

Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory & Cognition, 25,* 826-837.

McDermott, K. B. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin & Review, 4,* 582-586.

McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia, 41,* 293-303.

McDermott, K. B., & Watson, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory & Language, 45,* 160-176.

McKone, E., & Murphy, B. (2000). Implicit false memory: Effects of modality and multiple study presentations on long-lived semantic priming. *Journal of Memory & Language, 43,* 89-109.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning & Verbal Behavior, 16,* 519-533.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (Eds.), <u>Affect and accuracy in recall: Studies of "flashbulb memories</u>. New York: Cambridge University Press.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25,* 838-848.

Ochsner, K. N. (2000). Are affective events richly "remembered" or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General, 129,* 242-261.

Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally charged lures immune to false memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 328-338.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage, 16,* 331-348.

Raichle, M. E., Fiez, J. A., Videen, T. O., Macleod, A. M., Pardo, J. V., Fox, P. T. & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex, 4,* 8-26.

Rhodes, M. G., & Anastasi, J. S. (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin & Review, 7,* 158-162.

Richardson, M. P., Strange, B. A., & Dolan, R. J. (2004). Encoding of emotional memories depends on amygdala and hippocampus and their interactions. *Nature Neuroscience, 7,* 278-285.

Roediger, W. E. (1978). The aldehyde fuchsin and colloidal iron staining reactions in the canine thyroid C cell. *Histochem J, 10,* 549-57.

Roediger, H. L. III, (1990). Implicit memory: Retention without remembering. *American Psychologist, 45,* 1043-1056.

Roediger, H. L., III., Balota, D. A., & Watson, J. M. (2001). Spreading activation and the arousal of false memories. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), The nature of remembering: Essays in honor of Robert G. Crowder. Washington, DC: American Psychological Association. (pp. 95-115).

Roediger, H. L. III, & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word fragment completion. *Memory & Cognition, 15,* 379-388.

Roediger, H.L., III, Buckner, R. L., & McDermott, K. B. (1999). Components of processing. In J.K. Foster and M. Jelicic, Memory: Systems, Process, or Function? Oxford: Oxford University Press.

Roediger, H. L., III, & Gallo, D. A. (2001). Levels of processing: Some unanswered questions. In M. Naveh-Benjamin, M. Moscovitch, and H. L. Roediger (Eds.), Perspectives on human memory and cognitive aging: Essays in honour of Fergus Craik. New York, NY: Psychology Press.

Roediger, H. L., III, & McDermott, K. B. (2000). Distortions of memory. In E. Tulving and F.I.M. Craik (Eds.) Oxford Handbook of Memory. Oxford: Oxford University Press. P. 49-162.

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 803-814.

Roediger, H. L., III, & Payne, D. G. (1982). Hypermnesia: the role of repeated testing. *J Exp Psychol Learn Mem Cogn, 8,* 66-72.

Roediger, H. L., III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8,* 385-407.

Roediger, H. L., III, Weldon, M. S. & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. Varieties of memory and consciousness: Essays in honor of Endel Tulving. H. L. I. Roediger and F. I. M. Craik. Hillsdale, NJ, Erlbaum**:** 3-41.

Scarborough, D., Cortese, C., & Scarborough, H. (1977) Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1-17.

Schacter, D. L. (1994). Priming and multiple memory systems: Perceptual mechanisms of implicit memory. Memory Systems 1994. D. L. Schacter and E. Tulving. Cambridge, MA, MIT

Press**:** 244-256.

Schacter, D. L. (1990). Perceptual representation systems and implicit memory: Toward a resolution of the multiple memory systems debate. *Annals of the New York Academy of Sciences, 608,* 543-571.

Schacter, D. L. (1985). Priming of old and new knowledge in amnesic patients and normal subjects. *Annals of the New York Academy of Sciences, 444,* 44-53.

Schacter, D. L. & Buckner, R. L. (1998). Priming and the brain. *Neuron, 20,* 185-195.

Schacter, D.L., Dobbins, I.G., & Schnyer, D.M. (2004). Specificity of priming: A cognitive neuroscience perspective. *Nature Reviews Neuroscience*, *5,* 853-862.

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory & Language, 40,* 1-24.

Schacter, D. L., Reiman, E., Curran, T., Yun, L. S., Bandy, D., McDermott, K. B., & Roediger, H. L., III. (1996). Neuroanatomical correlates of veridical and illusory recognition memory: Evidence from positron emission tomography. *Neuron, 17,* 267-274.

Schacter, D.L. & Slotnick, S.D. (2004). The cognitive neuroscience of memory distortion. *Neuron, 44,* 149-160.

Schacter, D. L., Verfaellie, M., & Anes, M. D. (1997). Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology, 11*, 331-342.

Schacter, D.L., & Wiseman, A.L. (in press). Reducing memory errors: The distinctiveness heuristic. In R.R. Hunt & J. Worthen (Eds.), Distinctiveness and memory. New York: Oxford University Press.

Schmolck, H., Buffalo, E.A., & Squire, L.R. (2000). Memory distortions develop over time: recollections of the O.J. Simpson trial verdict after 15 and 32 months. *Psychological Science, 11,* 39-45.

Schnyer, D.M., Dobbins, I.G., Nicholls, L., Schacter, D.L., & Verfaellie, M. (in press). Rapid response learning in amnesia: Delineating associative learning components in repetition priming. *Neuropsychologia.*

Seamon, J. G., Goodkind, M. S., Dumey, A. D., Dick, E., Aufseeser, M. S., Strickland, S. E., Woulfin, J. R., & Fung, N. S. (2003). "If I didn't write it, why would I remember it?" Effects of encoding, attention, and practice on accurate and false memory. *Memory & Cognition, 31*, 445-457.

Seamon, J. G., Luo, C. R., Schwartz, M. A., Jones, K. J., Lee, D. M., & Jones, S. J. (2002). Repetition can have similar or different effects on accurate and false recognition. *Journal of Memory & Language, 46,* 323-340.

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience, 7,* 664-672.

Smith, R. E., & Hunt, R. R. (1998). Presentation modality affects false memory. *Psychonomic Bulletin & Review, 5,* 710-715.

Smith, S. M., Gerkens, D. R., Pierce, B. H., & Choi, H. (2002). The roles of associative responses at study and semantically guided recollection at test in false memory: The Kirkpatrick and Deese hypotheses. *Journal of Memory & Language, 47,* 436-447.

Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory & Language, 40,* 83-108.

Tenpenny, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review, 2,* 339-363.

Thapar, A., & McDermot, K. B. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition, 29,* 424-432.

Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory, 7,* 233-256.

Tulving, E. & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247,* 301-306.

Tulving, E., Schacter, D. L. & Stark, H. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 336-342.

Verfaellie, M., Page, K., Orlando, F., & Schacter, D. L. (in press). Impaired implicit memory for gist information in amnesia. *Neuropsychologia*

Warrington, E. K. & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia, 12,* 419-428.

Watson, J. M., McDermott, K. B., & Balota, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger-McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory & Cognition, 32,* 135-141.

Weaver, C. A., III. (1993). Do you need a "flash" to form a flashbulb memory? *Journal of Experimental Psychology: General, 122,* 39–46.

Weiss, A. P., Dodson, C. S., Goff, D. C., Schacter, D. L., & Heckers, S. (2002). Intact suppression of increased false recognition in schizophrenia. *American Journal of Psychiatry, 159,* 1506-1513.

Weldon, M. S. & Roediger, H. L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition, 15,* 269-280.

Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences, 97,* 11125-11129.

Whittlesea, B. W. A. (2002). False memory and the discrepancy-attribution hypothesis: The prototype-familiarity illusion. *Journal of Experimental Psychology: General, 131,* 96-115.

Wiggs, C. L. & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology, 8,* 227-33.

Zald, D.H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research: Brain Research Review, 41,* 88-123.

Acknowledgements

Figure 1

Figure xx. Recognition in the criterial-recollection task (Gallo, Weiss, & Schacter, 2004). On the Word Test, hits are to test words that were studied as red words, and false alarms are to test words that were studied as pictures (vice versa on the Picture Test). FAs to nonstudied items followed a similar pattern as FAs to to-be-excluded studied items.