

Functional Neuroimaging of False Memories

Nancy A. Dennis, Caitlin R. Bowman, and
Indira C. Turney

Introduction

Our memories are far from perfect. In fact we are prone to memory distortions that often render even the most vivid retrieval of past events inaccurate. One such example of the fallible nature of memory is that of *false memories*. A false memory refers to the situation in which we generate a memory of a past experience when in fact no such event occurred. Examples include remembering you took out the trash, when in fact you did not; remembering that you were told to pick up apples and bananas from the grocery store, when in fact oranges and pears were the to-be-purchased fruits; and believing you met a new acquaintance in one setting (e.g., a holiday party) when in fact you met him or her in another (e.g., a work seminar). Critical to the definition of false memory is that when making such an error of commission, an individual sincerely believes that the misremembered event actually occurred in the past. Often the false memory is not completely without precedence: on many occasions an event highly related to that which is falsely retrieved actually did occur; in other instances an individual may retrieve the correct event but misremember the source of the event (i.e., the time or place the event occurred). Given these similarities in phenomenology, true and false memories are often very difficult to differentiate.

Because of the overlap in the behavioral characteristics of true and false memories, and their high rate of occurrence in memory tests throughout the lifespan (McCabe *et al.*, 2009), the study of false memories represents a critical area of memory research. They have traditionally been examined using a variety of behavioral paradigms, but recent advances in neuroimaging have allowed for the use of positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and event-related potentials (ERPs) to elucidate the specific brain regions that support both true and false memories, as well as to identify a neural signature that differentiates the two types of memories. In this chapter we will briefly review findings from studies employing the most commonly utilized false memory paradigms in neuroimaging research. Within this review we will highlight neural processes that are common to true and false memories as well as those that distinguish between the two types of

memory processes. The review will include retrieval-related studies and encoding studies, and will conclude with a look at developmental studies of false memories.

False Memory Paradigms

While a wide variety of experimental paradigms have been developed to study false memories (for a review, see Brainerd and Reyna, 2005), only a small number have been adapted for use in neuroimaging. Those that have been so adapted focus on ensuring that the experimental paradigm produces a sufficient number of false memories in order to estimate a reliable neural signal. Accordingly, many false memory paradigms explicitly manipulate the relatedness between study and test items in order to produce ample false memories. The most widely used example is the Deese–Roediger–McDermott (DRM) paradigm (Deese, 1959; Roediger and McDermott, 1995). In the typical DRM paradigm, participants are presented with lists of words in which all the items in a given list are semantically related to an item that is not presented (the critical lure). For example, the words *bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, nap, peace, yawn,* and *drowsy* are all related to the critical related word, *sleep*. Though *sleep* is not presented in the original list of studied words, individuals tend to recall and/or recognize the critical lure at both a rate and confidence level similar to that associated with retrieval of the studied items (Roediger and McDermott, 1995). Due to the fact that the traditional DRM paradigm converges on one related word, thereby requiring long lists to be used in order to identify a single false memory, variations on the paradigm have been used in neuroimaging studies. One common variation includes using conceptual lists where a subset of the category exemplars are presented during study (e.g., farm animals: *chicken, sheep, pig, goat*) and several others are used as related lures at retrieval (e.g., *horse, cow*).

In addition to manipulating the semantic relatedness amongst items, many researchers have also used manipulations of perceptual relatedness to measure neural processes underlying false memories. Most neuroimaging studies utilizing this technique employ one of two major variations: categorized pictures (e.g., Gutchess and Schacter, 2012) or computer-generated abstract shapes (e.g., Slotnick and Schacter, 2004). In the categorical version of the paradigm, participants are presented with pictures of multiple exemplars from various categories during encoding (e.g., multiple exemplars of “chair”). During retrieval, target items are intermixed with related lures (category exemplars that were not presented at encoding) and unrelated lures (new items whose category was not presented during encoding). The use of computer-generated shapes works in a similar manner, whereby related lures are shapes that share perceptual overlap with a target while unrelated lures are more distinct.

A second major category of experimental paradigms designed to examine the neural basis of false memories is what we will label source-misattribution paradigms (i.e., imagery, misinformation, and source memory paradigms). In these paradigms, false memories are associated with incorrect memory for the original presentation, or source, of information. For example, in a typical imagery paradigm participants are provided with a list of items to study. While half of the items are accompanied by a pictorial representation of the item, the other half are presented only in word form

and the participants are asked to think about or imagine the item. At retrieval, memory is tested for which items were actually accompanied by a visual representation. False memories occur when participants believe that a presented word was accompanied by a visual representation, when in fact the participant had been asked to imagine the item. In the misinformation paradigm, participants are presented with an episode during an initial study phase (e.g., vignettes depicting a car accident). Following this initial presentation, information about the episode is presented again, with alterations (e.g., with a yield sign inserted in the vignette instead of the original stop sign). The new or altered information presented during this second phase is what is referred to as “misinformation.” At test, participants are asked to remember the original scenario of events. False memories occur when the misinformation is retrieved instead (e.g., remembering a yield sign instead of a stop sign). Finally, in source memory paradigms, false memories occur when, at retrieval, participants associate memory for one item with the source of a different item.

Neuroimaging of False Memories

Traditionally, false memories stemming from semantic or perceptual relatedness have been considered to be theoretically and qualitatively distinct from those stemming from source-misattribution paradigms (e.g., Brainerd and Reyna, 2005; Stark, Okado, and Loftus, 2010). In particular, related false memories are posited to occur due to retrieval of the shared semantic or perceptual gist across both targets and lures. On the other hand, false memories stemming from misattribution paradigms include information that has been explicitly presented or referenced during encoding. As such, imagery and misinformation paradigms often cite source confusion and misattribution (Johnson, Hashtroudi, and Lindsay, 1993) as the underlying mechanism supporting false memories and focus on the examination of neural processes during the study phase to help explain the occurrence of false memories (see also *Encoding studies* section, below). Despite differences in methodology and theoretical cause of false memories across both relatedness and source-misattribution paradigms, similar findings have been observed at retrieval.

Retrieval studies

By evoking a strong theme across encoded items, the DRM paradigm (and its variants) and perceptual false memory paradigms examine false memories associated with recognition responses to critical, or related, lures. Importantly, these critical lures share the same gist (i.e., semantic, conceptual meaning) or common features with items that were presented during encoding. As such, it is not surprising that the two paradigms exhibit highly similar findings, both with respect to commonalities in neural activity mediating true and false memories and with respect to neural correlates that differentiate based on the veracity of the memory.

Common neural processing One of the most ubiquitous findings generated from false memory studies is the considerable overlap in the neural networks mediating both true and false memories (Figure 8.1). Specifically, overlapping neural activity has

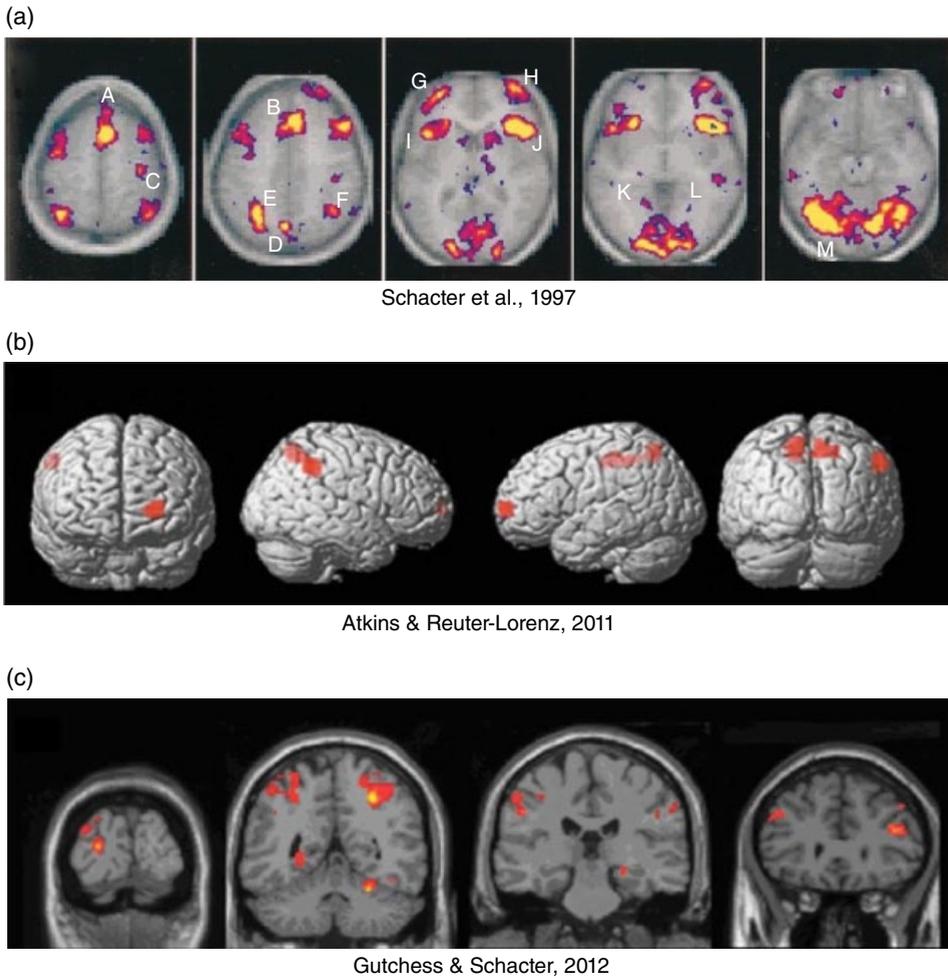


Figure 8.1 Results from relatedness paradigms showing common neural activity for true and false memories throughout frontal, parietal, temporal, and occipital cortices: (a) adapted from Schacter *et al.*, 1997; (b) adapted from Atkins and Reuter-Lorenz, 2011; (c) adapted from Gutchess and Schacter, 2012.

been found in bilateral frontal and parietal regions (Atkins and Reuter-Lorenz, 2011; Cabeza *et al.*, 2001; Dennis, Bowman, and Vandekar, 2012; Garoff-Eaton, Kensinger, and Schacter, 2007; Garoff-Eaton, Slotnick, and Schacter, 2006; Iidaka *et al.*, 2012; Kahn, Davachi, and Wagner, 2004; Schacter *et al.*, 1996, 1997; Slotnick and Schacter, 2004; Okado and Stark, 2003; von Zerssen *et al.*, 2001), bilateral caudate and insula (Stark, Okado, and Loftus, 2010; von Zerssen *et al.*, 2001), lateral temporal cortex (Cabeza *et al.*, 2001; Garoff-Eaton, Slotnick, and Schacter, 2006; Schacter *et al.*, 1996; Stark, Okado, and Loftus, 2010), occipital cortex (Dennis, Bowman, and Vandekar, 2012; Garoff-Eaton, Slotnick, and Schacter, 2006; Iidaka *et al.*, 2012; Schacter *et al.*, 1997; Slotnick and Schacter, 2004; Stark, Okado, and Loftus, 2010; von Zerssen *et al.*, 2001), and hippocampus/parahippocampal gyrus (PHG¹ Cabeza

et al., 2001; Dennis, Bowman, and Vandekar, 2012; Garoff-Eaton, Slotnick, and Schacter, 2006; Gutchess and Schacter, 2012; Kahn, Davachi, and Wagner, 2004; Schacter *et al.*, 1996, 1997; Slotnick and Schacter, 2004; Stark, Okado, and Loftus, 2010; von Zerssen *et al.*, 2001). This widespread overlap in neural activity mediating both true and false memories has been attributed to several factors, including the fact that targets and related lures share similar properties (e.g., Garoff-Eaton, Slotnick, and Schacter, 2006), the engagement of highly similar retrieval-related evaluation and monitoring processes (e.g., Atkins and Reuter-Lorenz, 2011), retrieval of contextual information (e.g., Okado and Stark, 2003), and evidence that both types of memories are supported by above-threshold familiarity processing (e.g., Kahn, Davachi, and Wagner, 2004).

While stimulus properties common to both targets and lures can include many features, in the case of relatedness paradigms these shared features often include a shared semantic meaning and/or perceptual similarity. By virtue of this similarity between targets and related lures, these stimuli are likely to evoke comparable retrieval-related activity within brain regions mediating both perceptual and semantic processing. Similar perceptual processing in late visual cortices (Brodmann areas [BA] 19 and 37) has been suggested to reflect processing that contributes to the conscious experience of memory, which is independent of true “oldness” (Slotnick and Schacter, 2004). Others have suggested that such processing is likely to reflect successful retrieval of the general properties of originally studied items such as shape and color (Garoff-Eaton, Slotnick, and Schacter, 2006) or those involving the semantic label or general category (e.g., fruit, bird) to which the item belongs. As such, common activity in semantic processing regions, such as left temporal gyrus and dorsolateral prefrontal cortex (PFC), has often been interpreted as retrieval of these semantic labels (Dennis, Kim, and Cabeza, 2008; Garoff-Eaton, Kensinger, and Schacter, 2007; Kim and Cabeza, 2007a; von Zerssen *et al.*, 2001).

Common activity in frontal regions has also been associated with retrieval effort and monitoring processes, presumed to operate independently of retrieval success (e.g., Atkins and Reuter-Lorenz, 2011). In this way, it is unsurprising that such activity underlies both true and false memory retrieval, as these processes support decision making, not necessarily the accuracy of a given decision or memory. Similarly, common activity in the precuneus and lateral parietal cortex is interpreted as reflecting general “recovery operations” (Cabeza *et al.*, 2001) or the general feeling of oldness (Atkins and Reuter-Lorenz, 2011) that also occurs independent of retrieval accuracy. Taken together, the results of studies that report common activity between true and false memories emphasize that activity across several different brain regions can reflect language, perceptual, and monitoring processes that are independent of the encoding history of the stimuli.

Distinct neural processing Despite finding such a large degree of overlap in neural recruitment, many studies also find differences in the level of neural engagement within regions mediating true and false memories. One of the most notable differences is that of increased activity in sensory cortices associated with true compared to false memories (Figure 8.2) (Abe *et al.*, 2008; Atkins and Reuter-Lorenz, 2011; Cabeza *et al.*, 2001; Dennis, Bowman, and Vandekar, 2012; Moritz *et al.*, 2006; Schacter *et al.*, 1996; Slotnick and Schacter, 2004). Greater engagement of sensory processing regions for true compared to false retrieval has most often been interpreted

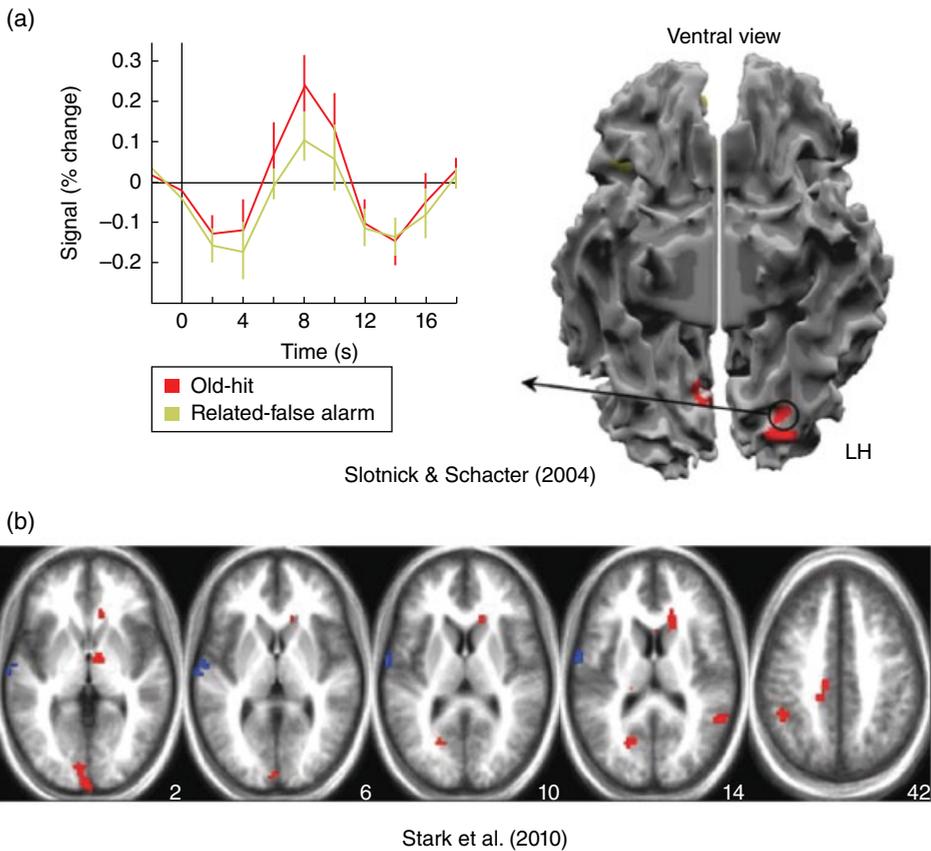


Figure 8.2 Activity supporting the sensory reactivation theory. (a) True recognition shows significantly greater activation in early visual cortex (BA 17/18) compared to false recognition (adapted from Slotnick and Schacter, 2004); (b) Regions showing a difference in activity for true versus false memory (red) include early visual cortical regions (BA 17/18) and striate cortex (adapted from Stark, Okado, and Loftus, 2010).

with respect to the “sensory reactivation hypothesis.” Founded in behavioral research that found true memories to be associated with more sensory and perceptual details than false memories (e.g., Marche, Brainerd, and Reyna, 2010; Mather, Henkel, and Johnson, 1997; Norman and Schacter, 1997), the theory posits that by virtue of having been presented previously, true memories will elicit reactivation of the encoding episode in sensory regions that were involved in their initial encoding – a finding that has been observed across several traditional memory studies (e.g., Vaidya *et al.*, 2002; Wheeler, Petersen, and Buckner, 2000). Having never been presented previously, false memories will not be accompanied by this heightened sensory signal. For example, in one of the earliest imaging studies of false memories, Schacter and colleagues (1996) utilized a verbal encoding task to present related words at study. While participants exhibited a highly similar network for both true and false memories, greater activity for true retrieval was found in the temporal–parietal junction, a region

associated with auditory processing. Researchers interpreted this finding as reflecting participants' retrieval of the auditory or phonological aspects of the item's presentation during the study phase. Several ERP studies have also observed differential activity over posterior cortices for true and false memories, supporting different amounts of sensory processing for true compared to false memories (Curran *et al.*, 2001; Fabiani, Stadler, and Wessels, 2000; Nessler and Mecklinger, 2003; Nessler, Mecklinger, and Penney, 2001; Walla *et al.*, 2000).

Recent findings from perceptual studies have both supported and expanded this earlier work, by showing not only increased activity in visual cortices for true memories, but also a dissociation between memory accuracy and recruitment of early versus late visual processing regions. Specifically, research has shown that activity in early visual processing regions (i.e., BA 17 and 18) distinguishes between true and false memories, while, as noted above, activity in late visual cortex (i.e., BA 19 and 37) is commonly active for true and false memories (Dennis, Bowman, and Vandekar, 2012; Slotnick and Schacter, 2004). While both early and late visual cortex are associated with object perception and identification, early visual cortex has been associated with recapitulation of a sensory signature (Buckner and Wheeler, 2001; Rugg and Wilding, 2000; Vaidya *et al.*, 2002) and late visual cortex has been linked to retrieval of general object identity and meaning (Vaidya *et al.*, 2002; Wheeler and Buckner, 2003; Wheeler, Petersen, and Buckner, 2000). Thus, with respect to true and false memories, common activity has been interpreted as reflecting retrieval of common perceptual details and conscious processing of an item as "old," whereas activity in early visual cortex has been interpreted as reflecting retrieval of perceptual and sensory details associated with the encoding episode (Dennis, Bowman, and Vandekar, 2012; Slotnick and Schacter, 2004; Stark, Okado, and Loftus, 2010).

Despite strong evidence supporting the sensory reactivation hypothesis, not all perceptual false memory studies find this dissociation (Garoff-Eaton, Slotnick, and Schacter, 2006; Gutchess and Schacter, 2012). For example, a recent study by Gutchess and Schacter (2012) found that as the gist representation was strengthened at encoding, the false-alarm rate increased, as did activation in both the hippocampus and in early and late visual processing regions (BA 17 and 37). The authors interpreted this increase in early and late visual cortex as indicative of a role for gist in sensory reactivation, suggesting that increased gist during false memories may reflect the retrieval of prototypical features shared by new and old items. Interestingly, high levels of gist had the opposite effect for true memories, such that increasing gist led to reduced visual activity, which the authors interpreted as reflecting the fact that true memories rely on parsing individual perceptual features and that this may be best supported under low gist conditions.

Consistent with findings from relatedness paradigms, several source-misattribution studies have found support for the sensory reactivation hypothesis, observing greater activity in sensory cortices for true compared to false retrieval (Fabiani, Stadler, and Wessels, 2000; Gonsalves *et al.*, 2004; Gonsalves and Paller, 2000; Kensinger and Schacter, 2006; Okado and Stark, 2003; Stark, Okado, and Loftus, 2010). For example, Okado and Stark (2003) also observed greater activity in several regions of visual cortices including right middle occipital cortex, left cuneus, left lingual gyrus, left fusiform gyrus for true compared to false memories. The authors concluded that this increase reflected the retrieval of item details associated with those items physically presented during encoding as opposed to items only imagined. More recently

the same group also showed support for the sensory reactivation hypothesis using a misinformation paradigm (Figure 8.2) (Stark, Okado, and Loftus, 2010). Specifically, researchers found that true memories of the original, visually presented event were accompanied by greater activation in early visual cortex (BA 17/18) compared to false memories of the auditorily presented misinformation. Interestingly, when the presentation of auditory misinformation led to a false recollection, participants' false retrieval was accompanied by greater activity in left superior temporal gyrus (i.e., auditory cortex, BA 22/42) than if the misinformation was correctly rejected. In both studies the authors attributed this increase in sensory cortex activity for true memories to the retrieval of sensory information associated with the original information presentation.

While the low spatial resolution afforded by ERP studies cannot differentiate between early and late visual processing, Gonsalves and Paller (2000) found that true memories exhibited greater processing in posterior sensory cortices, showing enhanced positivity in the 900–1200 ms window for true compared to false memories. The authors suggested that this difference may reflect the retrieval of visual details, which was greater for true memories and linked to the information encountered and stored at encoding. Fabiani and colleagues (2000) also saw differential activity across posterior electrode sites for true compared to false memories. In accord with the sensory reactivation theory, the authors concluded that increased positivity for true memories represented retrieval of the memory trace formed during encoding, whereas the absence of activity for false memories correctly indicated the absence of a sensory signature.

A second region that has been shown to differentiate true and false memories is the medial temporal lobe (MTL). While common activity in MTL regions has been identified (see above), several studies find that the MTL only supports retrieval of true memories or shows greater activity for true compared to false memories (Cabeza *et al.*, 2001; Dennis, Bowman, and Vandekar, 2012; Dennis, Kim, and Cabeza, 2008; Giovanello *et al.*, 2009; Kahn, Davachi, and Wagner, 2004; Kensinger and Schacter, 2006; Kim and Cabeza, 2007a; Paz-Alonso *et al.*, 2008). In several studies researchers have suggested that greater MTL activation for true memories reflects greater recovery of sensory details associated with true memories (Cabeza *et al.*, 2001; Kahn, Davachi, and Wagner, 2004; Okado and Stark, 2003), whereas others have suggested that this neural increase reflects the role of the hippocampus in binding together true details from past events (Kensinger and Schacter, 2006), or recollection processes (Dennis, Bowman, and Vandekar, 2012; Kim and Cabeza, 2007a). Considered with the fact that some studies find no MTL differences for true and false memories (Schacter *et al.*, 1997; Slotnick and Schacter, 2004; Stark, Okado, and Loftus, 2010), a satisfying theory regarding the role of the MTL in false memories may not be currently attainable, but several studies have tried to offer an explanation for the mixed results. For example, Cabeza and colleagues (2001) found that the anterior hippocampus exhibited similar activation for both true and false retrieval, whereas the posterior PHG showed greater activity for true retrieval (Figure 8.3). They suggested that activity in anterior regions reflects recovery of semantic information that supports both types of memories, whereas posterior PHG, by virtue of its connectivity with sensory cortices, reflects retrieval of sensory information specific to true memories.

Recently our lab found a similar mix of MTL results, observing common anterior hippocampus/PHG activity for both types of memories and greater right hippocampal

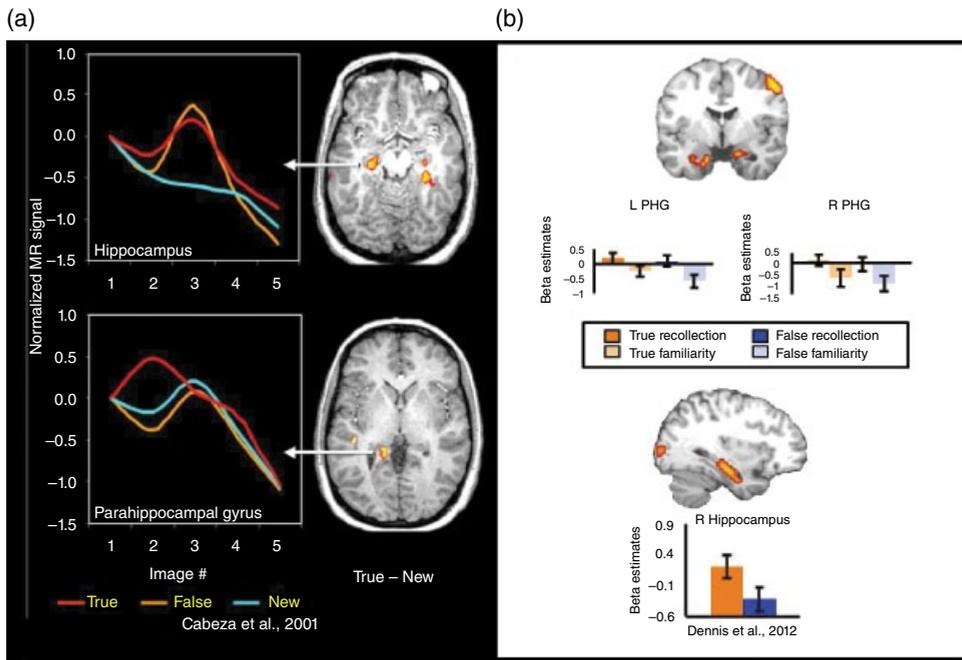


Figure 8.3 MTL activity during true and false memories. (a) The left anterior hippocampus shows common activity for both true and false compared to new items (upper panel), whereas the left posterior PHG shows increased activity only for true memories (lower panel) (adapted from Cabeza *et al.*, 2001; Copyright (2001) National Academy of Sciences of the USA. Reproduced with permission). (b) Activity in bilateral anterior PHG shows common activity for true and false recollection (upper panel), whereas activity in right hippocampus shows greater activity for true compared to false recollection (lower panel) (adapted from Dennis, Bowman, and Vandekar, 2012. Reproduced with permission of Elsevier).

activity for true compared to false recollection (Figure 8.3) (Dennis, Bowman, and Vandekar, 2012). Our results suggest that even though the MTL mediates retrieval processes leading to both true and false recollection, the hippocampus proper has the ability to distinguish between detailed recollection of true and erroneous events. We suggested that increased hippocampal activity may represent more details retrieved for true memories or the accurate reconstruction/binding of details supporting true recollection of past experiences. Furthermore, even though anterior MTL regions were commonly active in both true and false recollection, connectivity maps showed differential engagement of frontal, parietal, and occipital cortices for each type of memory. Specifically, true recollection was associated with functional connectivity within a more inferior network (including fusiform gyrus, hippocampus, middle temporal gyrus), whereas false recollection engaged a more superior network (including superior parietal, superior frontal gyrus, posterior cingulate cortex). We interpreted these differences as indicating that true recollection is driven by bottom-up integration of information from sensory input and item retrieval whereas false recollection is driven by top-down attention control processes. Thus, even though a given MTL region may mediate both types of memories, the processing involved may reflect subtle differences in the cognitive operations associated with true and false memories.

While differential activity in sensory cortices and the MTL has been the focus of analyses distinguishing true from false memories, increased activity in the prefrontal cortex has been shown to differentiate related false memories. Specifically, a number of studies have reported increased activity in bilateral prefrontal cortices for false compared to true memories (Cabeza *et al.*, 2001; Garoff-Eaton, Kensinger, and Schacter, 2007; Kensinger and Schacter, 2006; Kim and Cabeza, 2007a; Kubota *et al.*, 2006; Okado and Stark, 2003; Schacter *et al.*, 1996, 1997; Slotnick and Schacter, 2004). Studies have attributed increased PFC activity to monitoring, reconstructive processes, and semantic elaboration – with the interpretation dependent upon the specific locus of PFC activation. For example, in a DRM task, Schacter and colleagues (1996) found right dorsolateral/anterior PFC to be more active during false compared to true recognition. Given the role of this region in retrieval monitoring (e.g., Fletcher, Shallice, and Dolan, 1998; Henson, Shallice, and Dolan, 1999), the authors concluded that the activation may reflect the need for increased retrieval monitoring and evaluation associated with the strong familiarity evoked by the false memory. In an imagery study by Okado and Stark (2003), increased PFC activity for false memories was localized to the right anterior cingulate gyrus. Given the role of the anterior cingulate in response competition and conflict (Kerns *et al.*, 2004), the authors concluded that this reflects the increased effort involved in incorrectly endorsing an imagined item as “seen.” ERP studies also support the conclusion that frontal regions may distinguish between true and false memories, and be engaged in greater monitoring and evaluation associated with false retrieval (Curran *et al.*, 2001; Fabiani, Stadler, and Wessels, 2000; Goldmann *et al.*, 2003; Nessler, Mecklinger, and Penney, 2001; Wiese and Daum, 2006). For example, Nessler and Mecklinger (2003) observed that ERPs were more positive at frontal locations for false than true recognition across short retention delays (40s versus 80s). They suggested that this may reflect participants’ greater focus on related lures leading to false recollection in long delays compared to familiarity discrimination during short delays.

Others have interpreted increased activity in PFC as reflecting semantic elaboration underlying false memories (Cabeza *et al.*, 2001; Garoff-Eaton, Kensinger, and Schacter, 2007; Kubota *et al.*, 2006). For example, Garoff-Eaton and colleagues (2007) found increased activity in left inferior, middle, and medial frontal gyrus for false retrieval associated with lures that were conceptually related to encoded items (e.g., *silver, bronze*), but not when the lure was perceptually related to encoded items (e.g., *bell, tell*). Given that both types of false memories should require equivalent monitoring and evaluation, they attributed the increased left prefrontal activity to the retrieval of both conceptual information that pertained to item meaning and the semantic gist associated with studied items that also pertained to the related lures (see also evidence from encoding studies below). The idea that false memories are mediated by either familiarity or the gist representation associated with the encoding theme is a conclusion reached in several neuroimaging studies (Dennis, Kim, and Cabeza, 2008; Duarte, Graham, and Henson, 2010; Garoff-Eaton, Kensinger, and Schacter, 2007; Kim and Cabeza, 2007a, 2007b; Moritz *et al.*, 2006). This theory is supported not only by the above-mentioned increased false memory activity in left prefrontal cortices, but also by false memory activity in regions outside prefrontal cortex that are also associated with semantic processing. For example, in a semantic relatedness study, Moritz *et al.* (2006) found that, compared to true memories, false memories for critical lures were associated with activation in left inferior temporal lobe, a region

associated with general semantic processing (for a review, see Saumier and Chertkow, 2002; Wise and Price, 2006). The authors proposed that such activity may reflect semantic gist processing or the spreading of activation in semantic networks. Taken together, results suggest that when critical lures are semantically or perceptually related to study items, the semantic gist may evoke a sense of familiarity that is strong enough to form the basis of a false memory. Such a conclusion was posited by Kim and Cabeza (2007a), who found frontoparietal activity to mediate high-confidence false memories. Given the role of frontal and parietal cortex in familiarity (Cansino *et al.*, 2002; Yonelinas *et al.*, 2005), the authors concluded that a strong feeling of familiarity underlies false memories associated with critical lures that match the gist trace evoked during encoding (e.g., categorical information relating items). A similar familiarity argument is also made in support of unrelated lures as well (see Duarte, Graham, and Henson, 2010).

Encoding studies

While most studies examining the neural correlates of false memories have focused on the retrieval phase, it has been well argued that encoding processes also contribute to false memories. However, it is relatively difficult, from a methodological perspective, to design a study that isolates the neural processes that contribute to the formation of a false memory. For example, in the relatedness paradigms it is suggested that false memories arise from gist that is built up across many trials (see Brainerd and Reyna, 2002), and so activity on any given trial may only partially contribute to a false memory. Similarly, in misinformation paradigms false memories may arise due to processing during either the original or the misinformation phase. Thus, the practice of isolating a *single* time-point during encoding that would create a false memory is a difficult, and often unattainable, endeavor. Despite this challenge, a handful of studies have examined the influence of encoding processes on the formation of false memories (Baym and Gonsalves, 2010; Gonsalves *et al.*, 2004; Gonsalves and Paller, 2000; Kensinger and Schacter, 2005; Kim and Cabeza, 2007b; Kubota *et al.*, 2006; Okado and Stark, 2005) and have attempted to elucidate the cognitive and neural processes that underlie the formation of false memories.

Misinformation studies overcome the issue of localizing encoding-related false memory activity by defining true memories as those consistent with the original event phase, whereas false memories arise when information from the secondary event phase is remembered as if it were presented during the original event. Perhaps unsurprisingly, evidence from misinformation studies suggests that if an individual is presented with two sources of information, the information that receives the greater amount of neural processing (either original/true information or secondary/false information) is that which is most likely to be remembered (Baym and Gonsalves, 2010; Gonsalves *et al.*, 2004; Okado and Stark, 2005; Stark, Okado, and Loftus, 2010). For example, Okado and Stark (2005) found that encoding activity in the left hippocampus and perirhinal cortex was greater for true than for false memories in the original encoding phase, yet during the presentation of misinformation, activity in this region was greater for false than for true memories. The authors concluded that, given the role of the hippocampus in the encoding of source information (Davachi, Mitchell, and Wagner, 2003), greater contextual processing in this region during encoding reflected

which information would be ultimately bound to the encoding episode. Consistent with the above evidence, studies have also found that the degree of visual imagery engaged during encoding leads to subsequent false memories in reality-monitoring paradigms (Aminoff, Schacter, and Bar, 2008; Gonsalves *et al.*, 2004; Kensinger and Schacter, 2005). For example, Gonsalves and colleagues (2004) found that greater engagement of precuneus, right inferior parietal cortex, and anterior cingulate during imagery trials led to subsequent false memories (i.e., participants believing to have viewed the item during encoding) (see also Gonsalves and Paller, 2000, for ERP evidence). Citing the role of these regions in visual imagery tasks (e.g., Ishai, Ungerleider, and Haxby, 2000; Kosslyn and Thompson, 2000), the authors concluded that the enhanced visual imagery during encoding led to inaccurate memories. Taken together, these results suggest that false memories may be dependent on activation associated with both the presentation of the true information and the presentation of the misinformation episode.

The engagement of sensory cortices during encoding has also been investigated in relatedness studies. For example, in a modified DRM paradigm, Kim and Cabeza (2007b) used an encoding trial composed of four related items (e.g., chick, sheep, pig, goat) to examine encoding activity leading to subsequent true and false memories. By presenting all related items on a single trial, the authors attempted to capture that elusive single time-point when a false memory of related information would occur during study. They found that while regions involved in semantic elaboration (left ventro- and dorsomedial PFC) and conscious item processing (bilateral occipitotemporal and occipitoparietal cortex) were involved in both true and false memory formation. True memory formation showed greater activity in left posterior PHG and early visual cortex (BA 18/17) (Figure 8.4). Thus, results suggest that encoding of specific perceptual information supports subsequent true memories. Using a perceptual relatedness paradigm, Garoff-Eaton and colleagues (2006) found a dissociation in sensory processing for the encoding of false memories. Specifically, while right fusiform cortex was engaged for subsequent true memories, the left fusiform was engaged for both subsequent true memories and trials leading to the endorsement of similar lure items. Thus, the authors concluded that right fusiform supports encoding of visual details specific to an individual item, whereas left fusiform encodes more general perceptual information. The foregoing results suggest that false memory formation is a byproduct of elaborative semantic and visual processing, whereas the formation of true memories is also based on the encoding of sensory details and raw memory traces.

Overall, results suggest that encoding-related processing contributes to false memories in several ways. With respect to the encoding of item-specific details, and consistent with the sensory reactivation theory, findings suggest that encoding of detailed sensory information associated with true events supports subsequent veridical memories and lower occurrences of false memories. Secondly, greater processing of misinformation or greater engagement in imagery during the study phase will ultimately lead to a strong memory trace that is likely to be misidentified as part of the true memory in subsequent memory tests. With regard to related false memories, results suggest that processes which support true memories, such as elaborative semantic processing and encoding of gist, also form the foundation of subsequent false memories.

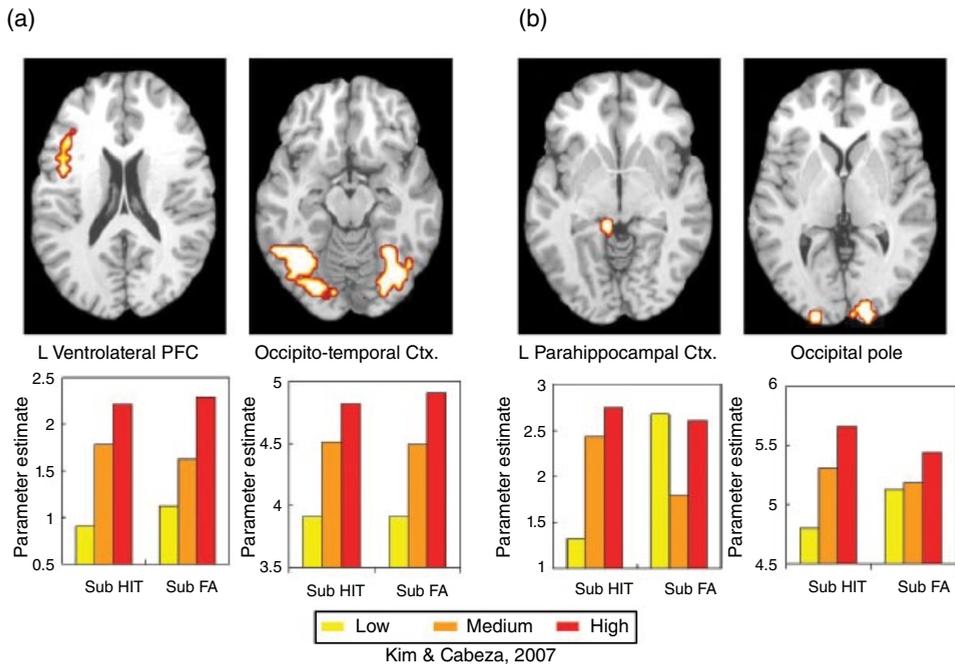


Figure 8.4 Common and distinct activity at encoding that predicts subsequent true and false memories. (a) Activity in left ventrolateral PFC and occipitotemporal cortex shows a similar pattern of activity for subsequent true and false memories. (b) Activity in left PHG and occipital pole shows increased activity for subsequent true memories, but not for subsequent false memories (adapted from Kim and Cabeza, 2007b. Reproduced with permission of Oxford University Press).

Unrelated false memory studies

While the foregoing results illustrate the neural mechanisms involved in related false memories and false memories that arise due to a targeted encoding manipulation, sometimes false memories arise that are unrelated to information presented at encoding. Often labeled as “unrelated false memories,” these false memories are interesting because there is seemingly no basis for the formation of the memory. Due to the difficult nature of obtaining a sufficient number of unrelated false memories for imaging analysis, very few neuroimaging studies have investigated the neural basis for these false memories (Duarte, Graham, and Henson, 2010; Garoff-Eaton, Slotnick, and Schacter, 2006; Iidaka *et al.*, 2012). Consequently, given the small number of studies and variety in methods and analyses used, there has been little consensus about the underlying neural process supporting such memories.

Garoff-Eaton and colleagues (2006) were the first to directly explore the neural basis of unrelated false memories and did so by using the abstract shape paradigm (described above). They found that the left superior and middle temporal gyri were more active for unrelated false recognitions than for either true or related false recognitions. The authors concluded that, given the role of these brain regions in mediating language processing, unrelated false recognition may arise as participants

misattribute a verbal tag generated at encoding to an unrelated item generated during retrieval. While a number of studies have shown that false memories arise from a reliance on verbal and semantic information as opposed to sensory information, it is unclear why the verbal labeling strategy would lead to unrelated and not related false recognitions, as it has in other studies (Cabeza *et al.*, 2001; Dennis, Kim, and Cabeza, 2007, 2008; Garoff-Eaton, Slotnick, and Schacter, 2006; Schacter *et al.*, 1996). Interestingly, no region exhibited overlap between unrelated and related false recognition and true recognition. The authors concluded that these two types of false memories do not share overlapping cognitive processes.

Unlike the study by Garoff-Eaton *et al.* (2006), a study by Duarte, Graham, and Henson (2010) found considerable overlap in the neural correlates mediating true and false memories, even when lures were unrelated line drawings. Specifically, overlap was seen in medial parietal, middle frontal, and lateral temporal regions. The authors attributed this activity to common familiarity processing shared by true and false recognition. One region that did distinguish between true and false memories was the anterior MTL. While the anterior MTL has been associated with common activity in previous studies (e.g., Cabeza *et al.*, 2001; Dennis, Bowman, and Vandekar, 2012), the authors suggested that this overlap was driven by semantic and/or perceptual similarities between old and new items that was not present amongst the unrelated items in their study. Given that the study did not include related lures, a conclusion cannot be drawn with respect to similarities between related and unrelated false memories.

Lastly, Iidaka and colleagues (2012) examined related and unrelated false memories of faces. They found that regions including left superior and inferior parietal, left inferior frontal gyrus, and early visual cortex (BA 18) supported unrelated false recognitions compared to unrelated correct rejections. However, several of these regions (including left superior parietal and left inferior frontal gyrus) were also active for related false recognition. Thus, without any direct comparisons or conjunction analyses, it is unclear the extent to which this activity is unique to unrelated false memory or is part of a network that reflects more general processes involved in false memories.

While it is difficult to draw general conclusions as to the neural basis of unrelated false memories from such a small sample of studies, data from both Duarte and colleagues (2010) and Iidaka and colleagues (2012) suggest that, like related false memories, the cognitive process mediating false memories to unrelated items may be similar to those underlying true memories. Specifically, results across all false memory paradigms suggest that the processes involved in memory retrieval are not based on the veridicality of the memory itself, but reflect the search and monitoring/evaluation process of stored representations.

Developmental studies

While the study of the neural basis of false memories is itself a relatively new endeavor, research is relatively limited with respect to developmental differences. Behavioral research has shown that developmental differences in false memories depend on the type of false memory being measured. For example, while both children and older adults are more susceptible to misinformation and source misattribution than are young adults, false memories stemming from semantic and perceptual relatedness

increase from childhood to adulthood, and throughout aging (see Brainerd and Reyna, 2005; Brainerd, Reyna, and Ceci, 2008). Though these age differences have been examined broadly at a behavioral level, only one study to date has examined the neural basis of false memories in children (Paz-Alonso *et al.*, 2008), and only a handful of neuroimaging studies have investigated the question in older adults (Dennis, Bowman, and Peterson, 2014; Dennis, Kim, and Cabeza, 2007, 2008; Duarte, Graham, and Henson, 2010; Giovanello *et al.*, 2009; Gutchess, Ieui, and Federmeier, 2007).

Using the DRM paradigm, Paz-Alonso and colleagues (2008) found significant developmental differences across frontal, parietal, and MTL regions mediating both true and false memories. Specifically, while the anterior MTL and the parietal cortex failed to distinguish between true and false memories in young children (8-year-olds), they showed a graded response distinguishing veridicality across older children (12-year-olds) and adults. With respect to false memories, researchers found that the left ventrolateral PFC exhibited increased activity for hits and related false memories in adults but not in children (ages 8 and 12). The authors suggested that this developmental change is indicative of more elaborate semantic processing with age. Similarly, evidence from aging studies has suggested that older adults (> 60 years old) exhibit a greater reliance on semantic processing than do younger adults, and it is proposed that this increase may underlie increases in semantic false memories across the lifespan (Dennis, Kim, and Cabeza, 2007, 2008).

With respect to false memories and aging, neuroimaging data has supported findings from behavioral studies showing age-related reductions in recollection processing associated with true memories and age-related reliance on gist and/or familiarity processing supporting both true and false memories (e.g., Balota *et al.*, 1999; Koutstaal and Schacter, 1997; Spencer and Raz, 1995; Tun *et al.*, 1998). Specifically, false memory studies have observed age deficits in neural activity mediating true recollection in both the MTL (Dennis, Kim, and Cabeza, 2007, 2008; Duarte, Graham, and Henson, 2010) and the visual cortex (Dennis, Bowman, and Peterson, 2014; Dennis, Kim, and Cabeza, 2007; Duarte, Graham, and Henson, 2010; see also Gutchess, Ieui, and Federmeier, 2007, for ERP evidence). In addition to overall decreases in visual activity, visual processing regions that differentiate true from false memories in young adults (i.e., fusiform gyrus, early visual cortex) have not exhibited the same differentiation in aging (Figure 8.5) (Dennis, Bowman, and Peterson, 2014; Duarte, Graham, and Henson, 2010; Gutchess, Ieui, and Federmeier, 2007). Taken together, these results support the theory that an age-related deficit in processing veridical information, and differentiating between studied and unstudied information within sensory cortices, contributes to age-related increases in false memories.

As noted, the prevailing finding in aging research is an age-related shift from recollection to familiarity and gist processing. False memory research has found that this shift in processing is associated with both true and false memories (Dennis, Bowman, and Peterson, 2014; Dennis, Kim, and Cabeza, 2007, 2008; Duarte, Graham, and Henson, 2010; Giovanello *et al.*, 2009). For example, using both semantic and perceptual related paradigms, Dennis and colleagues (Dennis, Bowman, and Peterson, 2014; Dennis, Kim, and Cabeza, 2007, 2008) have found both true and false memories to be mediated by the middle and superior temporal gyri, regions involved in semantic and gist processing (Saumier and Chertkow, 2002; Simons *et al.*, 2005). Giovanello and colleagues (2009) also found age-related increases in familiarity

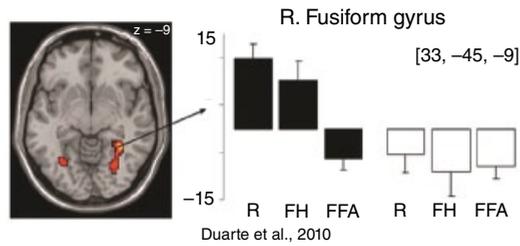


Figure 8.5 Age-related dedifferentiation in visual processing regions. Young adults (black bars) show increased activity in the right fusiform gyrus for true recollection (R) and true familiarity (FH) compared to familiar false alarms (FFA) while older adults (white bars) do not show differentiation between true and false memories (adapted from Duarte, Graham, and Henson, 2010).

processing within right PHG supporting false memories. Duarte, Graham, and Henson (2010) found that both recollection and familiarity-related activity actually showed evidence of age-related deficits, concluding that the similarities in familiarity processing and reduced differentiation between true and false memories led to increased false memories in older adults. Taken together, the results suggest that older adults encode information in a less detailed fashion and thus do not have those details available to them at retrieval, thereby making the representation of true and false memories more similar.

Conclusions and Future Directions

Overall, neuroimaging studies have shown that, by and large, the neural correlates that mediate retrieval of false memories overlap with those supporting true memory retrieval. Taking into account the commonalities that often exist between targets and related lures, as well as the reconstructive nature of memory retrieval, this is not a surprising finding. Specifically, the results suggest that retrieval of common object properties such as semantic label or context elicit similar processing across true and false memories. In addition, the same evaluation and monitoring processes that are involved in identifying true memories are also active when retrieving false memories. Despite such significant overlap, neural differences with respect to the veridicality of memories do emerge. These differences are most commonly reported as greater MTL and visual cortex activity for true relative to false memories, and greater activity in frontal cortices for false relative to true memories. Greater activity for true compared to false memories in visual cortices (and other sensory processing regions) has been interpreted with respect to the sensory reactivation hypothesis. The idea that true memories are accompanied by the retrieval of greater sensory details compared to false memories may also underlie the finding across several studies that MTL activity is greater for true compared to false memories. Several interpretations regarding increased PFC activity for false memories have also been offered, including a greater reliance on gist and familiarity processes as well as greater monitoring associated with making what may be a more difficult memory decision.

Unfortunately, despite the observed differences between true and false memories, no neural signal or pattern of activation has shown itself to be a reliable neural marker of either true or false memory across all studies. Sensory reactivation, while replicated across several studies, is not ubiquitous. Nor is increased frontal activity supporting false memories. While differences in methodologies and stimuli can account for some variations in the data, these discrepancies require further investigation. Future research should focus not only on replications of current findings, controlling both for stimuli content and methodological differences, but also on analyzing patterns of brain activity and use of effective connectivity methods. One promising line of research, specifically with regard to elucidating the role of the MTL in false memories, involves high-resolution fMRI and multi-voxel pattern analysis (see Chapter 6 for more discussion of these techniques). In particular, recent work by Yassa and colleagues (Lacy *et al.*, 2011; Yassa and Stark, 2008, 2011; Yassa *et al.*, 2011) suggests that different subregions within the MTL respond differentially to true and false memories, possible accounting for the mix of results observed using traditional fMRI methods.

In addition to understanding the neural basis of false memories during retrieval studies, it is widely accepted that encoding processes also contribute to the occurrence of false memories. However, pinpointing the neural processes that lead to subsequent false memories has been challenging. With respect to relatedness studies, results have suggested that subsequent false memories are associated with familiarity and gist processing during encoding, whereas false memories stemming from imagery and misinformation are associated with enhanced processing of the misinformation as compared to details associated with the true encoding event. To this end, encoding evidence also supports the sensory reactivation theory, finding that subsequent true memories are associated with greater initial sensory processing, and that greater engagement of sensory processing during encoding reduces subsequent false memories.

A third important finding from neuroimaging research has been that fact that the neural correlates mediating false memories differ across the lifespan. Mirroring previously observed age-related differences in episodic memory (see Chapters 17 and 18), false memory studies also find age-related deficits in neural processing mediating true memories, including deficits in sensory-related processing during both encoding and retrieval. Age-related deficits in processing details of episodic memories may lead to a degraded representation of the event, thereby making an individual vulnerable to false memories. Supporting the idea of dedifferentiation in aging, true as well as false memories in older adults have been shown to be mediated by both gist and familiarity processing.

While the current research represents a solid start to understanding the neural basis of false memories, more research is needed. For example, despite finding increased frontal activity for false memories, research to date has not identified a biomarker or neural signal clearly differentiating false from true memories. In addition, while a number of studies have found that true memories elicit greater activity in sensory regions, possibly reflecting reactivation of sensory details from study, the exact location of this activity has differed from study to study, and no study has directly tested the reactivation theory, comparing encoding and retrieval activity. Finally, despite the large number of neuroimaging studies, very few studies have directly linked their results to theories of false memories – e.g., source monitoring theory (Johnson, Hashtroudi, and Lindsay, 1993), activation monitoring theory (Roediger

and McDermott, 1995), and fuzzy trace theory (Brainerd and Reyna, 2002). In the same vein, the foregoing theories of false memory were originally developed to account for behavioral data, and thus do not include hypotheses and predictions regarding neural activations. Future neuroimaging research should do more than simply report results, but should attempt to couch its findings in terms of behavioral theories. Moreover, behavioral theories should be expanded to generate testable neural predictions.

Acknowledgments

This research was supported by a National Science Foundation (NSF) Grant BCS1025709 awarded to N.A.D. and was conducted while N.A.D. was an AFAR Research Grant recipient from the American Federation for Aging Research. We wish to thank Avery Rizio and Kristina Peterson for helpful comments through the writing process.

Note

- 1 This may underestimate the overlap in medial temporal lobe (MTL) regions, as the specific analyses reported across several studies precluded us from drawing definitive conclusions.

References

- Abe, N., Okuda, J., Suzuki, M., *et al.* (2008). Neural correlates of true memory, false memory, and deception. *Cerebral Cortex*, 18 (12), 2811–2819.
- Aminoff, E., Schacter, D.L., and Bar, M. (2008). The cortical underpinnings of context-based memory distortion. *Journal of Cognitive Neuroscience*, 20 (12), 2226–2237. doi: 10.1162/jocn.2008.20156.
- Atkins, A.S., and Reuter-Lorenz, P.A. (2011). Neural mechanisms of semantic interference and false recognition in short-term memory. *NeuroImage*, 56 (3), 1726–1734. doi: 10.1016/j.neuroimage.2011.02.048.
- Balota, D.A., Cortese, M.J., Duchek, J.M., *et al.* (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's Type. *Cognitive Neuropsychology*, 16, 361–384.
- Baym, C.L., and Gonsalves, B.D. (2010). Comparison of neural activity that leads to true memories, false memories, and forgetting: an fMRI study of the misinformation effect. *Cognitive, Affective and Behavioral Neuroscience*, 10 (3), 339–348. doi: 10.3758/CABN.10.3.339.
- Brainerd, C.J., and Reyna, V.F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, 11, 164–169.
- Brainerd, C.J., and Reyna, V.F. (2005). *The Science of False Memory*. Oxford Psychological Series. Oxford: Oxford University Press.
- Brainerd, C.J., Reyna, V.F., and Ceci, S.J. (2008). Developmental reversals in false memory: a review of data and theory. *Psychological Bulletin*, 134 (3), 343–382.
- Buckner, R.L., and Wheeler, M.E. (2001). The cognitive neuroscience of remembering. *Nature Reviews Neuroscience*, 2 (9), 624–634.

- Cabeza, R., Rao, S.M., Wagner, A.D., *et al.* (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 of the USA*, 98 (8), 4805–4810.
- Cansino, S., Maquet, P., Dolan, R.J., and Rugg, M.D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cerebral Cortex*, 12 (10), 1048–1056.
- Curran, T., Schacter, D.L., Johnson, M.K., and Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience*, 13 (2), 201–216.
- Davachi, L., Mitchell, J.P., and Wagner, A.D. (2003). Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences of the USA*, 100 (4), 2157–2162.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58 (1), 17–22.
- Dennis, N.A., Bowman, C.R., and Peterson, K.P. (2014). Age-related differences in the neural correlates mediating false recollection. *Neurobiology of Aging*, 35 (2), 395–407. doi: 10.1016/j.neurobiolaging.2013.08.019.
- Dennis, N.A., Bowman, C.R., and Vandekar, S.N. (2012). True and phantom recollection: an fMRI investigation of similar and distinct neural correlates and connectivity. *NeuroImage*, 59 (3), 2982–2993. doi: 10.1016/j.neuroimage.2011.09.079.
- Dennis, N.A., Kim, H., and Cabeza, R. (2007). Effects of aging on the neural correlates of true and false memory formation. *Neuropsychologia*, 45, 3157–3166.
- Dennis, N.A., Kim, H., and Cabeza, R. (2008). Age-related differences in brain activity during true and false memory retrieval. *Journal of Cognitive Neuroscience*, 20 (8), 1390–1402.
- Duarte, A., Graham, K.S., and Henson, R.N. (2010). Age-related changes in neural activity associated with familiarity, recollection and false recognition. *Neurobiology of Aging*, 31 (10), 1814–1830.
- Fabiani, M., Stadler, M.A., and Wessels, P.M. (2000). True but not false memories produce a sensory signature in human lateralized brain potentials. *Journal of Cognitive Neuroscience*, 12 (6), 941–949.
- Fletcher, P.C., Shallice, T., and Dolan, R.J. (1998). The functional roles of prefrontal cortex in episodic memory. I. Encoding. *Brain*, 121 (7), 1239–1248.
- Garoff-Eaton, R.J., Kensinger, E.A., and Schacter, D.L. (2007). The neural correlates of conceptual and perceptual false recognition. *Learning and Memory*, 14 (10), 684–692.
- Garoff-Eaton, R.J., Slotnick, S.D., and Schacter, D.L. (2006). Not all false memories are created equal: the neural basis of false recognition. *Cerebral Cortex*, 16 (11), 1645–1652.
- Giovanello, K.S., Kensinger, E.A., Wong, A.T., and Schacter, D.L. (2009). Age-related neural changes during memory conjunction errors. *Journal of Cognitive Neuroscience*, 22 (7), 1348–1361.
- Goldmann, R.E., Sullivan, A.L., Droller, D.B., *et al.* (2003). Late frontal brain potentials distinguish true and false recognition. *Neuroreport*, 14 (13), 1717–1720. doi: 10.1097/01.wnr.0000087908.78892.23.
- Gonsalves, B., and Paller, K.A. (2000). Neural events that underlie remembering something that never happened. *Nature Neuroscience*, 3 (12), 1316–1321.
- Gonsalves, B., Reber, P.J., Gitelman, D.R., *et al.* (2004). Neural evidence that vivid imagining can lead to false remembering. *Psychological Science*, 15 (10), 655–660.
- Gutchess, A.H., Ieuji, Y., and Federmeier, K.D. (2007). Event-related potentials reveal age differences in the encoding and recognition of scenes. *Journal of Cognitive Neuroscience*, 19 (7), 1089–1103. doi: 10.1162/jocn.2007.19.7.1089.
- Gutchess, A.H., and Schacter, D.L. (2012). The neural correlates of gist-based true and false recognition. *NeuroImage*, 59 (4), 3418–3426. doi: 10.1016/j.neuroimage.2011.11.078.

- Henson, R.N., Shallice, T., and Dolan, R.J. (1999). Right prefrontal cortex and episodic memory retrieval: a functional MRI test of the monitoring hypothesis. *Brain*, 122 (7), 1367–1381.
- Iidaka, T., Harada, T., Kawaguchi, J., and Sadato, N. (2012). Neuroanatomical substrates involved in true and false memories for face. *NeuroImage*, 62 (1), 167–176. doi: 10.1016/j.neuroimage.2012.04.044.
- Ishai, A., Ungerleider, L.G., and Haxby, J.V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, 28 (3), 979–990.
- Johnson, M.K., Hashtroudi, S., and Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin*, 114 (1), 3–28.
- Kahn, I., Davachi, L., and Wagner, A.D. (2004). Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. *Journal of Neuroscience*, 24 (17), 4172–4180.
- Kensinger, E.A., and Schacter, D.L. (2005). Emotional content and reality-monitoring ability: fMRI evidence for the influences of encoding processes. *Neuropsychologia*, 43 (10), 1429–1443.
- Kensinger, E.A., and Schacter, D.L. (2006). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *Journal of Neuroscience*, 26 (9), 2564–2570.
- Kerns, J.G., Cohen, J.D., MacDonald, A.W., III, et al. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303 (5660), 1023–1026.
- Kim, H.K., and Cabeza, R. (2007a). Trusting our memories: dissociating the neural correlates of confidence in veridical and illusory memories. *Journal of Neuroscience*, 27, 12190–12197.
- Kim, H.K., and Cabeza, R. (2007b). Differential contributions of prefrontal, medial temporal, and sensory-perceptual regions to true and false memory formation. *Cerebral Cortex*, 17, 2143–2150.
- Kosslyn, S.M., and Thompson, W.L. (2000). Neural systems activated during visual mental imagery. In *Brain Mapping: The Systems* (ed. A. W. Toga and J. C. Mazziotta). San Diego, CA: Academic Press, pp. 535–560.
- Koutstaal, W., and Schacter, D.L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, 37, 555–583.
- Kubota, Y., Toichi, M., Shimizu, M., et al. (2006). Prefrontal hemodynamic activity predicts false memory: a near-infrared spectroscopy study. *NeuroImage*, 31 (4), 1783–1789. doi: 10.1016/j.neuroimage.2006.02.003.
- Lacy, J.W., Yassa, M.A., Stark, S.M., et al. (2011). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learning and Memory*, 18 (1), 15–18. doi: 10.1101/lm.1971111.
- Marche, T.A., Brainerd, C.J., and Reyna, V.F. (2010). Distinguishing true from false memories in forensic contexts: can phenomenology tell us what is real. *Applied Cognitive Psychology*, 24 (8), 1168–1182.
- Mather, M., Henkel, L.A., and Johnson, M.K. (1997). Evaluating characteristics of false memories: remember/know judgments and memory characteristics questionnaire compared. *Memory and Cognition*, 25 (6), 826–837.
- McCabe, D.P., Roediger, H.L., McDaniel, M.A., and Balota, D.A. (2009). Aging reduces veridical remembering but increases false remembering: neuropsychological test correlates of remember-know judgments. *Neuropsychologia*, 47 (11), 2164–2173. doi: 10.1016/j.neuropsychologia.2008.11.025.
- Moritz, S., Glascher, J., Sommer, T., et al. (2006). Neural correlates of memory confidence. *NeuroImage*, 33 (4), 1188–1193. doi: 10.1016/j.neuroimage.2006.08.003.
- Nessler, D., and Mecklinger, A. (2003). ERP correlates of true and false recognition after different retention delays: stimulus- and response-related processes. *Psychophysiology*, 40 (1), 146–159.

- Nessler, D., Mecklinger, A., and Penney, T.B. (2001). Event related brain potentials and illusory memories: the effects of differential encoding. *Cognitive Brain Research*, 10 (3), 283–301.
- Norman, K.A., and Schacter, D.L. (1997). False recognition in younger and older adults: exploring the characteristics of illusory memories. *Memory and Cognition*, 25 (6), 838–848.
- Okado, Y., and Stark, C. (2003). Neural processing associated with true and false memory retrieval. *Cognitive, Affective and Behavioural Neuroscience*, 3 (4), 323–334.
- Okado, Y., and Stark, C.E. (2005). Neural activity during encoding predicts false memories created by misinformation. *Learning and Memory*, 12 (1), 3–11. doi: 10.1101/lm.87605.
- Paz-Alonso, P.M., Ghetti, S., Donohue, S.E., et al. (2008). Neurodevelopmental correlates of true and false recognition. *Cerebral Cortex*, 18 (9), 2208–2216.
- Roediger, H.L., and McDermott, K.B. (1995). Creating false memories: remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 8033–8014.
- Rugg, M.D., and Wilding, E.L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences*, 4 (3), 108–115.
- Saumier, D., and Chertkow, H. (2002). Semantic memory. *Current Neurology and Neuroscience Reports*, 2 (6), 516–522.
- Schacter, D.L., Buckner, R.L., Koutstaal, W., et al. (1997). Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *NeuroImage*, 6 (4), 259–269.
- Schacter, D.L., Reiman, E., Curran, T., et al. (1996). Neuroanatomical correlates of veridical and illusory recognition memory: evidence from positron emission tomography. *Neuron*, 17 (2), 267–274.
- Simons, J.S., Verfaellie, M., Hodges, J.R., et al. (2005). Failing to get the gist: reduced false recognition of semantic associates in semantic dementia. *Neuropsychology*, 19 (3), 353–361.
- Slotnick, S.D., and Schacter, D.L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, 7 (6), 664–672.
- Spencer, W.D., and Raz, N. (1995). Differential effects of aging on memory for content and context: a meta-analysis. *Psychology and Aging*, 10 (4), 527–539.
- Stark, C.E., Okado, Y., and Loftus, E.F. (2010). Imaging the reconstruction of true and false memories using sensory reactivation and the misinformation paradigms. *Learning and Memory*, 17 (10), 485–488. doi: 10.1101/lm.1845710.
- Tun, P.A., Wingfield, A., Rosen, M.J., and Blanchard, L. (1998). Response latencies for false memories: gist-based processes in normal aging. *Psychology and Aging*, 13 (2), 230–241.
- Vaidya, C.J., Zhao, M., Desmond, J.E., and Gabrieli, J.D. (2002). Evidence for cortical encoding specificity in episodic memory: memory-induced re-activation of picture processing areas. *Neuropsychologia*, 40 (12), 2136–2143.
- von Zerssen, G.C., Mecklinger, A., Opitz, B., and von Cramon, D.Y. (2001). Conscious recollection and illusory recognition: an event-related fMRI study. *European Journal of Neuroscience*, 13 (11), 2148–2156.
- Walla, P., Endl, W., Lindinger, G., et al. (2000). False recognition in a verbal memory task: an event-related potential study. *Cognitive Brain Research*, 9 (1), 41–44.
- Wheeler, M.E., and Buckner, R.L. (2003). Functional dissociation among components of remembering: control, perceived oldness, and content. *Journal of Neuroscience*, 23 (9), 3869–3880.
- Wheeler, M.E., Petersen, S.E., and Buckner, R.L. (2000). Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences of the USA*, 97 (20), 11125–11129.

- Wiese, H., and Daum, I. (2006). Frontal positivity discriminates true from false recognition. *Brain Research*, 1075 (1), 183–192. doi: 10.1016/j.brainres.2005.12.117.
- Wise, R.J.S., and Price, C.J. (2006). Functional imaging of language. In *Handbook of Functional Neuroimaging of Cognition*, 2nd edn (ed. R. Cabeza and A. Kingstone). Cambridge, MA: MIT Press, pp. 191–228.
- Yassa, M.A., Lacy, J.W., Stark, S.M., *et al.* (2011). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus*, 21 (9), 968–979. doi: 10.1002/hipo.20808.
- Yassa, M.A., and Stark, C.E. (2008). Multiple signals of recognition memory in the medial temporal lobe. *Hippocampus*, 18 (9), 945–954. doi: 10.1002/hipo.20452.
- Yassa, M.A., and Stark, C.E.L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, 34 (10), 515–525. doi: 10.1016/j.tins.2011.06.006.
- Yonelinas, A.P., Otten, L.J., Shaw, K.N., and Rugg, M.D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, 25 (11), 3002–3008.