RECOGNITION MEMORY IN AMNESTIC MILD COGNITIVE IMPAIRMENT

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Neuroscience

By

Lauren Elizabeth Ullrich, M.S.

Washington, DC
April 22, 2014
RECOGNITION MEMORY IN AMNESTIC MILD COGNITIVE IMPAIRMENT

Lauren Elizabeth Ullrich, M.S.

Thesis Advisor: Rhonda B. Friedman, Ph.D. and R. Scott Turner, M.D., Ph.D.

ABSTRACT

A detailed characterization of the memory impairment in amnestic mild cognitive impairment (aMCI) is important both clinically and in the interests of research. The progressive nature of dementia and the fact that neuronal loss often precedes behavioral symptoms means that early detection of Alzheimer’s disease will be crucial for any preventative strategy. In addition, because memory loss occurs in tandem with medial temporal lobe degeneration, studies of this population can provide insight into the nature of the brain structures underlying recognition memory. The current study investigated memory impairment in aMCI through a dual-process framework, which holds that recognition memory is supported by two processes: recollection and familiarity. While familiarity has generally been found to be preserved in healthy aging, there is some evidence that this process may be impaired in aMCI, though reports have been inconsistent. In the current study, participants with aMCI and age-and education-matched controls were tested on the remember-know task, the confidence judgment procedure, and the process dissociation procedure. On all three tasks, aMCI participants demonstrated impairment in both recollection and familiarity. Additionally, these processes were correlated with the volumes of four MTL structures: the hippocampus and the entorhinal, perirhinal, and posterior parahippocampal cortices. Consistent with the dual-process model, across aMCI and control participants, the strongest relationships were between recollection and hippocampal volume and familiarity and entorhinal cortex volume. The results from this study suggest that measures of
familiarity, in particular, may have utility as a proxy for AD neuropathology, but the predictive value of this measure still requires additional study.
I dedicate this dissertation to my mom, dad, and sister—without their unwavering confidence and support I never would have finished. I can’t list all the ways in which they have assisted me—thank you.

I also want to thank:

My advisors, Rhonda Friedman and Scott Turner, for their mentorship, time, and assistance;

My funding sources, the Cosmos Club, Sigma Xi, the MCGSO Student Research Grant Program, the Georgetown Biomedical Graduate Program, the Neural Injury and Plasticity NIH training grant, and the IPN NIH training grant;

My thesis committee, Darlene Howard, Michael Ullman, Jean Wrathall, and Howard Eichenbaum for their advice and encouragement;

David Wolk and Bradford Dickerson for their generous guidance, especially during the early stages of this project;

John VanMeter and CFMI for help collecting scans and providing development time on the scanner;

Patrick Forcelli and Ludise Malkova for their thoughtful comments on this manuscript;

The Friedman lab, past and present, including my fellow graduate students and compatriots, Elizabeth Lacey and Monika Mellem; our amazing SLPs, Sarah Snider, Aimee Carney, and Susan Lott; and our indefatigable research assistants, especially Tamar Gefen and Leah Orchinik—you are all amazing, strong, beautiful women and it has been my privilege to work (and play) with you;

My matriculating class, especially Patrick Forcelli, Elizabeth West Niedringhaus, and Emily Waterhouse, for their confidence and camaraderie, for being up to talk about anything, scientific or silly, and for being there to share the highs and lows of graduate school;

The entire Interdisciplinary Program in Neuroscience, students and faculty, for their community and support. Over these past years, I have come to realize what a unique program we have, and greatly appreciate the work and dedication that goes into sustaining it.

Many thanks,

LAUREN E. ULLRICH
TABLE OF CONTENTS

Introduction............................................................................................................................................... 1

1.1 Alzheimer’s disease and mild cognitive impairment ................................................................. 1

1.2 Recognition memory...................................................................................................................... 3

1.2.1 Dual-process theory of recognition memory ........................................................................ 3

1.2.1.1 Familiarity ............................................................................................................................ 4

1.2.1.2 Recollection ................................................................................................................................ 5

1.2.2 Models of recognition memory ............................................................................................ 5

1.2.2.1 The dual-process signal detection model .................................................................................. 6

1.2.2.2 The unequal-variance signal detection model ........................................................................ 7

1.2.2.3 Other models of recognition memory .................................................................................... 9

1.3 Separating recollection and familiarity .................................................................................... 10

1.3.1 Task dissociation ..................................................................................................................... 10

1.3.2 Process estimation ................................................................................................................... 11

1.3.2.1 Remember-know paradigm ................................................................................................... 12

1.3.2.2 Confidence judgment procedure .......................................................................................... 13

1.3.2.3 Process dissociation procedure ............................................................................................ 15

1.3.3 Conclusion ................................................................................................................................ 16

1.4 The medial temporal lobe and memory ...................................................................................... 16

1.4.1 Medial temporal lobe anatomy .................................................................................................. 17

1.4.1.1 Perirhinal cortex ....................................................................................................................... 18

1.4.1.2 Parahippocampal cortex ......................................................................................................... 18

1.4.1.3 Entorhinal cortex .................................................................................................................... 19

1.4.1.4 Hippocampus ........................................................................................................................... 20

1.4.2 Incorporating anatomy into the models .................................................................................. 21

1.4.2.1 Dual-process models................................................................................................................ 21

1.4.2.2 Other models .......................................................................................................................... 22

1.4.3 Evidence for an anatomical distinction .................................................................................. 24

1.4.3.1 Functional imaging ................................................................................................................. 24

1.4.3.2 Lesion studies ......................................................................................................................... 26

1.4.3.3 Recollection and familiarity in animals .................................................................................. 29

1.4.3.4 Conclusion ............................................................................................................................... 33
LIST OF ABBREVIATIONS

AD – Alzheimer’s disease

ADLs/IADLs – activities of daily living/instrumental activities of daily living

ApoE – apolipoprotein E

BIC – binding items and contexts

CS – collateral sulcus

DPSD – dual-process signal detection

ERC – entorhinal cortex

HC - hippocampus

LEC – lateral entorhinal cortex

MCI/aMCI/naMCI – mild cognitive impairment/ amnestic MCI/ non-amnestic MCI

MEC – medial entorhinal cortex

MTL – medial temporal lobe

PDP – process-dissociation procedure

PPHC – parahippocampal cortex

PRC – perirhinal cortex

R-K task – remember-know task

ROC/zROC – receiver operating characteristic/z-transformed ROC

ROI – region of interest

UVSD – unequal-variance signal detection
INTRODUCTION

1.1 Alzheimer’s disease and mild cognitive impairment

Alzheimer’s disease (AD) is the most common cause of dementia, affecting roughly 20% of adults over the age of 75 (Hebert et al., 2003). In 2012, an estimated 5.4 million Americans had Alzheimer’s disease (Alzheimer’s Association 2012), and this number is only expected to increase as baby boomers age (Terry Jr. et al., 2011). Patients with AD have impairment in memory and at least one additional cognitive domain; meet the criteria for dementia (cognitive impairment that interferes with usual functioning); and these symptoms had a gradual onset and continue to worsen over time (McKhann et al., 2011).

Mild cognitive impairment (MCI) is considered to be the prodromal stage of dementia. Interest in MCI has increased as it has become clear that effective treatment and/or prevention of AD will depend on identifying the disease before the onset of dementia because by the time of AD diagnosis, neuropathology is too far advanced for effective treatment (Terry Jr. et al., 2011). Neuropathological evidence of AD is present years before clinical symptoms occur (Morris et al., 2001; Price and Morris, 1999), and is more common in non-demented individuals with memory impairment than in those without memory impairment (DeCarli, 2003).

MCI is defined as impairments in one or more cognitive functions without significant deterioration in activities of daily living (ADLs; Petersen, 2004; Winblad et al., 2004). Patients with MCI typically convert to dementia (particularly Alzheimer’s disease), at a much higher rate than age-matched controls—around 5 to 15% per year (Farias et al., 2009).
To increase diagnostic specificity, MCI has been subdivided into two different types defined by the nature of the cognitive impairment—specifically by the presence or absence of memory deficits, namely amnestic MCI (aMCI) versus non-amnestic MCI (naMCI). aMCI has a higher likelihood of progression to AD than naMCI, which is more likely to progress to other forms of dementia, such as primary progressive aphasia or frontotemporal dementia (Busse et al., 2006; Mariani et al., 2007; Tabert et al., 2006; Yaffe et al., 2006). Both of these types can also be characterized by impairment in a single cognitive domain (e.g., memory alone) or multiple domains (e.g., memory and executive functioning; Petersen, 2004).

The consensus criteria for aMCI require subjective worsening memory, objective memory impairment, normal cognitive function in other domains, and minimal impairment in instrumental ALDs (Petersen, 2004; Winblad et al., 2004). Prevalence of aMCI ranges from 3% to 12% in the United States (Ward et al., 2012), with multi-domain aMCI being more prevalent—estimates ranging from 0.5 to 16% compared to estimates of the prevalence of single domain aMCI ranging from 0.5 to 8% (Jak et al., 2009).

Though there are not specific batteries of cognitive tests or even established cut-off scores to identify aMCI, diagnosis is often made on the basis of neuropsychological tests such as the Logical Memory subtest of the Wechsler Memory Scale, which measures delayed recall of a passage (Mariani et al., 2007); impairment is usually defined as performance worse than 1.5 standard deviations below age- and education-matched controls. Depending on the patient’s diagnosis (such as multi-domain aMCI), other areas of deficit may also include spatial recall and navigation or executive functioning (Jak et al., 2009). Beyond standard neuropsychological tests,
there has been somewhat limited investigation into the nature of the memory impairment in aMCI.

A detailed characterization of the memory impairment in aMCI is important both clinically and in the interests of research. Because of the progressive nature of dementia and the fact that neuronal loss often precedes behavioral symptoms, early detection will be crucial in any preventative strategy. Understanding how memory breaks down in early AD may facilitate the development of new behavioral interventions or drug therapies (Ally et al., 2009). It may also aid in refining the criteria for aMCI and reducing heterogeneity of research populations (Wolk et al., 2008).

This study investigates the nature of the memory impairment in aMCI using an approach informed by theories developed in the cognitive neuropsychology of declarative memory. In addition, this study explores the relationship between the volumes of brain areas implicated in declarative memory and various measures of memory impairment in an attempt to further inform our anatomical understanding of these processes.

1.2 Recognition memory

The hallmark deficit of aMCI is in declarative memory, which canonically refers to the ability to consciously recall both facts and past events (though it also includes unconscious processes). One widely studied example of this type of memory is recognition memory, which is the ability to judge whether or not a stimulus has been encountered before.

1.2.1 Dual-process theory of recognition memory

The most widely accepted theory of recognition memory, the so-called dual-process theory,
postulates that recognition memory is supported by two independent mechanisms: familiarity and recollection (Jacoby, 1991; Jacoby and Dallas, 1981; Mandler, 1980; Yonelinas, 2002; Yonelinas and Jacoby, 1994). Familiarity refers to the sense of “knowing” that you have seen an object before, in the absence of any associative information. Recollection refers to being able to retrieve contextual information about an episode in which the object was previously encountered.

The classic example of these two processes is the “butcher on the bus” phenomenon, in which a known person (the butcher) is encountered in a novel environment (the bus), leading to a prominent sense of familiarity without recollection of associative information (“I am sure I know that person, but from where?”). Subsequent recall of the identity of the person brings a host of associative information (her name, where she works, etc.), demonstrating recollection.

Although recollection and familiarity normally work together to support recognition, there is substantial evidence that they are two distinct memory processes with different characteristics and are functionally dissociable (Yonelinas, 2002).

### 1.2.1.1 Familiarity

Familiarity is thought to be a fast, predominantly automatic and unconscious process. As suggested by the definition, it is also an item-specific process, though the exact nature of the representations supported by familiarity is the subject of some debate. Most dual process theories contrast item familiarity with “associative” recollection (Brown and Aggleton, 2001; Yonelinas, 2002), in which subjects must associate two or more items, but some theorists argue that inter-item associative recognition can also be supported by familiarity (Mayes et al., 2007).

It has been demonstrated that certain forms of associative information are processed by the
familiarity system if they are “unitized” into one item, such as two words combined into one compound word (e.g., “sea” and “horse” processed as “seahorse;” Diana et al., 2008). However, Montaldi and Mayes (2010) argue that even non-unitized associative recognition as probed in recombination recognition tests can be supported by familiarity in certain cases (Bastin et al., 2010; Harlow et al., 2010). These issues, while important, are of less relevance to the traditional recognition memory tests used in this study, which only probe memory for single items.

Historically, the term “familiarity” has been used only for item recognition. “Familiarity” for places or contextual information may also exist, but this process has not been well studied. Here, according to convention, “familiarity” refers to item recognition.

1.2.1.2 Recollection

Recollection establishes links between pieces of information, such as an item and the context it was encountered in. Recollection is a more “qualitative” process wherein subjects can recall many aspects of a prior encounter, such as when and where it occurred, as well as any unique associations with that particular stimulus. In contrast to familiarity, recollection is a slower and more effortful process.

1.2.2 Models of recognition memory

Before these insights from cognitive neuropsychology can be applied to memory in aMCI, the theoretical framework with which to investigate these processes must be chosen. Although the general principle of the dual-process theory is generally accepted, there remains a debate over which of several models best fit the data. The two models that have been the most researched are the dual process signal detection model and the unequal variance signal detection model.
1.2.2.1 The dual-process signal detection model

The most widely studied instantiation of the dual-process theory is the dual-process signal detection (DPSD) model as proposed and modeled by Yonelinas (2002). This model makes specific predictions about the nature of these two processes.

In the DPSD model, familiarity is characterized by a graded process (comprised of weak to strong memories) that can support varying levels of confidence. It is assumed to be well-described by signal-detection theory because the memory strength distributions of old and new items form overlapping Gaussian distributions (see Figure 1-1). It is represented as $d'$, i.e., the difference in strength between previously encountered (“old”) and new items.

Recollection is described as a thresholded (i.e., it can succeed or fail), high-confidence process.

Both recollection and familiarity are able to support simple item recognition judgments, but only recollection can support judgments of source memory (where the item was previously encountered). The DSPD model makes no predictions about the underlying shape of the recollection distribution, only that it may succeed or fail, and is thus well represented by a simple probability. Recollection may also be a graded process—the quality and amount of information that subjects may recollect will differ—but an item will only be recognized if its memory strength exceeds a threshold (for discussion, see Parks and Yonelinas, 2007; Yonelinas et al., 2010).

The DPSD model has been elaborated and expanded into the Binding Items and Contexts (BIC) model to include explicit predictions about the different neural substrates of these processes in
the medial temporal lobe (Eichenbaum et al., 2007), as described in the section below,

\textit{Incorporating anatomy into the models.}

1.2.2.2 The unequal-variance signal detection model

The most well-described memory strength model, the unequal-variance signal detection (UVSD) model, proposes that recognition memory performance can be accounted for by a single, strength-based variable (Donaldson, 1996; Squire et al., 2007; Wixted, 2007). According to this model, tasks that purport to separate familiarity and recollection are actually differentiating weak and strong memories, respectively, on a continuum of memory strength.

In the UVSD model, recognition decisions are based on the strength of a graded memory signal (Figure 1-1). An item that has a memory strength greater than the decision criterion is declared old; otherwise it is new (Wixted, 2007). The model is so named because the variability of the memory strength distribution of old items is thought to be greater than for new items (i.e., unequal; Wixted, 2007). This is the result of differing amounts of memory strength added to items during the learning phase. Some items are learned well and memory strength increases greatly, while other items are not well learned and the increase in memory strength is slight; thus, the variability of the memory strength distribution of the old items is greater than the variability of the distribution of the new items. Like the DPSD model, it contains two parameters: a measure of memory strength and a measure of the variance of distribution of memory strength (equivalent to the standard deviation of the distribution).
It is important to note that, despite their critiques of the tasks used to measure recollection and familiarity, proponents of the UVSD model don’t necessarily deny the existence of two processes. There is now considerable evidence that pure memory-strength models are unable to account for performance on every type of recognition task (Diana et al., 2006; Wixted, 2007; Yonelinas, 2002), so the two processes have been incorporated into the UVSD model and other memory-strength models. The dual-process view can be reconciled with the UVSD model by considering “memory strength” as representing the sum of familiarity and recollection, which are combined prior to a recognition decision (Kelley and Wixted, 2001; Rotello et al., 2004; Wixted and Stretch, 2004). In this view, a decision about an individual item may be the result of a combination of both sets of information, regardless of the strength of the memory. Thus, the UVSD model is not necessarily a single process model, though the different contributions of the two processes are not modeled.
1.2.2.3 Other models of recognition memory

Dual-process theories of recognition memory similar to the DPSD have been put forth by other investigators (Onyper et al., 2010; Rotello et al., 2004; Sherman et al., 2003; Wixted and Mickes, 2010). Most of the dual process models are in agreement about the central features of recollection and familiarity, though they do differ in some specific predictions. None of these alternative models, however, have received the same attention as the UVSD model.

Most of these models can describe the data well, especially for standard recognition memory tasks—described later—as used in this study. It is important to consider that the main goal of the DPSD is to “develop a relatively simple model that is useful for applied problems such as the diagnosis and treatment of memory deficits (e.g., Jacoby et al., 1996), as well as for basic research in a variety of domains,” as opposed to developing a complex model that can completely describe performance on every task (Yonelinas and Jacoby, 2012). By this standard, most of the models described in the literature would suffice. The DPSD has particular utility because it provides quantitative estimates of recollection and familiarity and there is a large body of literature to provide comparative data on other populations (e.g., young adults, older adults, and individuals with amnesia) and conditions.

However, other theorists have argued that recollection and familiarity do not capture the relevant distinctions between the functions of various medial temporal lobe structures that underlie recognition memory, and thus dual process models do not have utility for understanding memory. This possibility will be discussed in more detail in The medial temporal lobe and memory.
1.3  **Separating recollection and familiarity**

By their very nature, recollection and familiarity usually work together to support recognition memory; thus, sophisticated methods of estimating their strength were developed to tease apart the contribution of each to recognition performance. Historically, this has been accomplished in two ways: by task dissociation and process estimation. Each of these approaches, while individually subject to caveats, when considered as a whole have largely supported the dual-process theory of recognition memory.

1.3.1  **Task dissociation**

Task dissociation procedures attempt to create tasks that isolate one or both of the processes underlying recognition memory by utilizing unique aspects of recollection and familiarity (e.g., the fact that recollection entails retrieval of qualitative information about the item). In these studies, recall performance (thought to reflect recollection) is often compared with recognition performance (thought to reflect the contribution of either or both processes). These studies do not attempt to quantify recollection and familiarity, so it is difficult to directly compare the strength of the processes or magnitude of impairment in memory-compromised populations.

Behavioral studies employing speeded responses have consistently suggested the existence of two processes by showing that familiarity is available earlier than recollection. When subjects were merely required to discriminate between studied and non-studied items, forcing them to speed their responses had little effect on accuracy. However, when subjects were forced to make speeded responses on associative tasks, which require remembering the item plus some aspect of
the context it was encountered in, performance suffered significantly (Gronlund et al., 1997; Hintzman and Caulton, 1997; Yonelinas and Jacoby, 1994).

Similarly, when subjects are required to discriminate between items that are highly familiar (a task for which familiarity alone is uninformative), the probability of a false alarm decreases as the time available to respond is increased (Dosher, 1984; Dosher and Rosedale, 1991). These results also suggest that recollective information is available later than familiarity. Single process models have difficulty accounting for these results (Diana et al., 2006).

1.3.2 Process estimation

The alternative approach to task dissociation is process estimation, which produces quantitative measures of recollection and familiarity strength from performance on a single task. Process estimation tasks include the remember-know task (R-K; Tulving, 1985), the confidence judgment procedure (Yonelinas and Jacoby, 1994); and the process-dissociation procedure (PDP; Jacoby, 1991).

In general, process-estimation tasks involve interpreting observed performance with a set of model equations to derive the parameter estimates representing the differential contributions of recollection and familiarity to performance (Yonelinas, 2002). Despite the differing assumptions made in each task, results from the different tasks have been remarkably consistent (Yonelinas, 2001). The UVSD model has also been used to interpret the results from many of these paradigms.
1.3.2.1 Remember-know paradigm

The remember-know (R-K) paradigm takes advantage of the subjective experiences of the two memory processes. In this paradigm, subjects are tested on a list of items, half of which have been seen previously. They rate whether the item is “new” or “old;” then, if they chose “old,” whether the object is “remembered” (they recall qualitative information about the study event) or “known” (the object is familiar). According to the dual-process model, recollection is mapped onto “remembering” a previously seen stimulus, while familiarity is mapped onto “knowing” you have seen the stimulus before.

Critics of dual-process models interpret the R-K paradigm as measuring differences in confidence as opposed to two different memory processes. In this view, a large reduction in remember responses is the expected result of a weakened memory trace; that is, fewer “remember” responses represent fewer high-confidence responses (Donaldson, 1996).

However, while the two responses subjectively map onto the experience of the two processes, there is also evidence that they are theoretically consistent with the DPSD model equation. Three lines of evidence are consistent with the idea that “remember” responses represent a high-confidence, thresholded process that can support associative information (recollection): high-confidence remember responses, accurate associative memory, and functionally dissociable responses.

Comparisons of R-K responses with confidence responses support the DPSD model (Gardiner and Java, 1991; Parkin and Walter, 1992; Rajaram, 1993). Under strict instructions to make a remember response only when the subject could, if asked, report what he or she recollected about
the item, remember responses were associated only with the highest level of confidence (Yonelinas, 2001; Yonelinas et al., 1996). Under less strict instructions, remember responses were associated with a wider range of confidence judgments (Rotello et al., 2005; Stretch and Wixted, 1998). The effect of instruction was replicated within one study (Rotello et al., 2005), suggesting that if “remember” is not defined correctly, subjects will make a remember response on the basis of high levels of familiarity.

In addition, “remembered” items are associated with accurate associative memory judgments that subjects are unable to produce for “known” items (Perfect et al., 1996; although see: Wais et al., 2008), which suggests that they differ in the types of mnemonic information they represent, not just in strength.

Finally, variables such as deep encoding (versus shallow encoding) and maintenance rehearsal increase one response and not the other (i.e., remember and know, respectively; Gardiner, 1988; Gardiner and Java, 1990, 1991; Gardiner et al., 1994; Rajaram, 1993, 1996), indicating that they are functionally dissociable.

1.3.2.2 Confidence judgment procedure

In the confidence judgment paradigm, subjects are tested on a list of items, half of which they have seen previously. In the test phase, subjects are asked to rate their confidence that they have seen the item before on a scale from 1 (certain it was not studied) to 6 (certain it was studied). From these responses, a receiver operating characteristic (ROC) is generated (Figure 1-2). In ROCs, hits (old items identified as old) are plotted against false alarms (new items identified as old) as a function of confidence.
In the dual-process account, the ROC is quantified by fitting a function to the data from which estimates of recollection and familiarity can be generated. In this model, the y-intercept represents recollection, because it supports high-confidence responses (plotted as the left-most point on the graph), while the degree of curvilinearity of the ROC represents familiarity, because it is proposed to increase gradually over a range of confidences (Eichenbaum et al., 2007). Tests of these assumptions have generally supported them (Yonelinas, 2001).

The UVSD model can also be fit to ROC data. In this model, each item is associated with a particular memory strength reflecting the degree of certainty that the item has been seen before. An item will be identified as “old” only if it exceeds some criterion of memory strength (Wixted, 2007).

In many cases both the DPSD and the UVSD models can provide a good fit for the ROC data from recognition memory tests. One area where they contain differential predictions is the shape...
of the z-transformed ROC curve (zROC). The DSPD model predicts that the zROC will be slightly U-shaped, representing the contribution of recollection to high-confidence recognition responses (Yonelinas and Jacoby, 1994). In contrast, the UVSD model predicts a linear zROC with a slope less than one, which reflects the ratio of the standard deviations of the distributions of the new and old items. In general, the zROC has been found to be U-shaped, especially with regards to relational recognition tasks thought to require recollection (Yonelinas and Parks, 2007). However, these two possibilities have been historically difficult to separate out in standard item recognition tasks (such as used in this study), as the U-shape predicted by the DPSD model is so slight under these conditions (Parks and Yonelinas, 2007).

1.3.2.3 Process dissociation procedure

The process dissociation procedure (PDP) pits recollection and familiarity against each other (Jacoby, 1991). The rationale behind the PDP is that recollection of an item will support determinations about the study event (such as the font color it was presented in), whereas familiarity will not (Yonelinas, 2002). Estimates for each process are generated by contrasting performance when both processes act together with performance when the two processes act in opposition (Yonelinas and Jacoby, 1995).

The strength of this approach is that recollection requires memory of associative or featural information encoded during the study event, making the PDP an objective measure of recollection, as opposed to a subjective measure like the R-K procedure (this aspect may be particularly attractive with respect to populations with brain damage, who may have reduced conscious insight into the quality or strength of their memories). Its strength is also a potential
weakness; the process-dissociation procedure uses a restricted measure of recollection.

Recollection in these tasks must be able to support a *specific* piece of associative information (e.g., what color font the word was presented in), as opposed to any other aspect of the study event (e.g., that it was the first item presented), and thus may underestimate recollection (Yonelinas and Jacoby, 1996). Critics of the dual process theory propose that the PDP task is merely separating strong, associatively-rich memories from weak memories.

### 1.3.3 Conclusion

Each task used to separate out the contributions of recollection and familiarity to recognition memory performance has assumptions and potential confounds that influence the interpretation of the results. The remember-know task and confidence judgment tasks depend on subjects’ insight into the quality of their memories while recollection in the PDP only measures memory for specific associative information. In addition, all three of these tasks assume that recollection and familiarity are independent. Task-dissociation procedures do not make the independence assumption, but also cannot provide quantitative measures of recollection and familiarity.

Thus, convergent evidence across paradigms is important in interpreting the data. The current study employed all three of the process estimation tasks described above to investigate recollection and familiarity in aMCI, with the expectation that the three tasks will provide converging evidence of deficits in these processes.

### 1.4 The medial temporal lobe and memory

Since patient H.M. (Scoville and Milner, 1957), evidence has shown the importance of the medial temporal lobe (MTL) for recognition memory. Decades of data from amnestic patients
with MTL lesions show that they have profound recognition memory deficits (Yonelinas et al., 1998), as do primates (Zola et al., 2000; Zola-Morgan et al., 1992), and rats (Fortin et al., 2004) with similar lesions. These findings are of particular importance to aMCI because the neuroanatomical hallmark of aMCI is medial temporal lobe (MTL) degeneration (Whitwell et al., 2007). Insight into the functions of these structures can aid in understanding of this and other diseases.

1.4.1 Medial temporal lobe anatomy

The MTL is broadly divided into four structures: the hippocampus and the parahippocampal (sometimes termed “posterior parahippocampal” in humans and postrhinal in rodents), perirhinal, and entorhinal cortices (together referred to as the parahippocampal area). These areas are roughly homologous across mammalian species, but their specific names and borders have only recently been standardized (for example, previous work in animals did not always make a distinction between perirhinal and postrhinal cortex; for discussion of these issues, see Burwell et al., 1995; Suzuki and Amaral, 2003).

The hippocampus is reciprocally connected to all the subregions of the parahippocampal area, and there are extensive interconnections between cortical areas as well (Suzuki and Amaral, 2003). Anatomical and electrophysiological evidence suggests that information flows from the cortical areas to the hippocampus in a network of parallel pathways that work together to support learning and memory (Burwell, 2000; Witter et al., 2000a, 2000b).
1.4.1.1 Perirhinal cortex

The perirhinal cortex (PRC) mainly receives input from association areas that process unimodal sensory information about objects, including olfactory, auditory, somatosensory and visual information (Burke et al., 2012; Suzuki, 1996; Suzuki and Amaral, 2003). These areas are part of the ventral visual processing stream, or “what” stream (Goodale and Milner, 1992; Mishkin and Ungerleider, 1982). The perirhinal cortex is also heavily interconnected with both the parahippocampal cortex and lateral entorhinal cortex, and receives input from the hippocampal areas CA1 and subiculum (Kealy and Commins, 2011; Suzuki, 1996). The perirhinal cortex also has prominent reciprocal connections with the amygdala (Pitkänen et al., 2000; Suzuki, 1996).

From the neuroanatomical connections, a picture emerges of the perirhinal cortex as a “zone of convergence” for sensory information from association areas as well as information from subcortical structures. The perirhinal cortex is in a position to synthesize multi-modal sensory information and interact with other MTL structures involved in memory (Suzuki, 1996). The prominent inputs from the “what” stream also suggest a special involvement in object recognition memory. Several groups have argued that the perirhinal cortex is also involved in the perception of complex stimuli by “unitizing” elements into one entity (Bussey and Saksida, 2005; Bussey et al., 2002; Kent and Brown, 2012; Murray and Wise, 2012; Murray et al., 2007).

1.4.1.2 Parahippocampal cortex

The posterior parahippocampal (postrhinal in the rat) cortex (PPHC) receives substantial input from areas associated with the dorsal visual processing stream, or “where” pathway (Suzuki and Amaral, 2004). The PPHC has reciprocal connections with both perirhinal cortex and medial
entorhinal cortex (Burwell and Amaral, 1998a, 1998b). It is much less well connected to the amygdala than the perirhinal cortex (Stefanacci et al., 1996).

In contrast to the involvement of the perirhinal cortex in object recognition, these anatomical connections of the parahippocampal cortex suggest that it may play a larger role in spatial memory, synthesizing information in the spatial domain and interacting with other memory structures. This has been extrapolated in some theories to suggest that the parahippocampal cortex plays a more general role in processing “contextual” information (Diana et al., 2006), although it must be noted that the boundaries of this term are poorly defined and there has not been much systematic investigation of what properties bias stimuli or groups of stimuli to be processed as “context” rather than individual objects (Montaldi and Mayes, 2010). In humans, much attention has been paid to the “parahippocampal place area,” a region in posterior parahippocampal cortex that responds preferentially to visual scenes (Epstein and Kanwisher, 1998; Epstein and Ward, 2010).

1.4.1.3 Entorhinal cortex

The entorhinal cortex (ERC) receives 60% of its input from the perirhinal and parahippocampal cortex (Insausti et al., 1987a, 1987b). Perirhinal cortex projects most heavily to the lateral entorhinal cortex and PPHC to the medial entorhinal cortex (Suzuki and Amaral, 1994). These projections mostly remain segregated in the entorhinal cortex (Eichenbaum et al., 2007). The strongest cortical input to entorhinal cortex can be characterized as high-level polymodal association cortex (Mohedano-Moriano et al., 2007). Interestingly, the direct cortical input to the lateral and medial entorhinal cortex (LEC and MEC) parallels the direct input to the perirhinal
cortex and PPHC, respectively (Burwell, 2000). This separation of inputs to the entorhinal cortex is modest, but suggests a parallel organization that persists through the different levels of the MTL.

The entorhinal cortex is reciprocally connected with all MTL structures. There is evidence for two functionally distinct parallel routes between the entorhinal cortex and hippocampus, with the projections of the LEC and MEC to the hippocampus remaining relatively separate in both the rat and the macaque (Canto et al., 2008; Deshmukh and Knierim, 2011; Eichenbaum and Lipton, 2008; Hargreaves et al., 2005). Convergence of the two streams largely occurs in the hippocampus (Eichenbaum and Lipton, 2008).

1.4.1.4 Hippocampus

The main source of cortical input to the hippocampus (HC) is the entorhinal cortex through the perforant pathway (Witter et al., 2000a). The medial and lateral entorhinal cortex both send projections to the entire HC. These projections converge on the same neurons in the dentate gyrus and CA3, but separate neurons in the subiculum and CA1 (Witter et al., 2000a). It is theorized that this organization may underlie the ability of the hippocampus to associate events and their context as well as distinguish between them. The hippocampus also receives direct projections from the perirhinal cortex and PPHC as well as other cortical regions (Naber et al., 1999).

The cortical outputs of hippocampal processing, arising in CA1 and subiculum, involve feedback connections back to the parahippocampal area, which projects back to the neocortical association areas from which the inputs to the medial temporal area originated. This organization suggests a
central role in organizing or extending the persistence of cortical representations (Eichenbaum, 2000).

1.4.2 Incorporating anatomy into the models

1.4.2.1 Dual-process models

Based on the anatomy described previously, proponents of the dual process model propose an anatomical dissociation within the MTL in which various MTL structures support different aspects of recognition memory.

One way this model has been articulated is the Binding Items and Contexts (BIC) model (Diana et al., 2007; Eichenbaum et al., 1992, 2007; Ranganath, 2010a, 2010b). This model proposes that the hippocampus and PPHC support recollection and the perirhinal and entorhinal (particularly lateral entorhinal) cortices support familiarity (Aggleton and Brown, 1999; Brown and Aggleton, 2001; Eichenbaum et al., 2007). In addition, extra-MTL structures such as prefrontal cortex are thought to support recollection, while the amygdala supports familiarity.

According to this model, item information from the “what” pathway is processed by the perirhinal cortex and lateral entorhinal cortex. Contextual information from the “where” stream is processed by the PPHC and medial entorhinal cortex. These two streams converge in the hippocampus, which “binds” object information in its context (Eichenbaum et al., 2012; Figure 1-3). More specifically, the BIC model proposes that activation of an item representation in perirhinal cortex (familiarity) is forwarded to the hippocampus, which may cause pattern completion of the activity elicited during the learning event and feedback to activate contextual
information in the PPHC and associated cortical areas, leading to recollection (Ranganath, 2010b).

**Figure 1-3. Binding Items and Contexts model.**
Based on Eichenbaum and colleagues (2007), the BIC model proposes that perceptual processing of object features by neocortical areas (the “what” stream) converges on the perirhinal cortex and lateral entorhinal cortex, while contextual (e.g. spatial) information from other neocortical areas (the “where” stream) converges on the posterior parahippocampal cortex (PPHC) and medial entorhinal cortex. These streams remain relatively segregated until they converge on the hippocampus. Information also flows back through reverse projections.

Similar models proposed by other groups (Aggleton and Brown, 1999; Mayes et al., 2007; McClelland et al., 1995; Montaldi and Mayes, 2010; Norman and O’Reilly, 2003; O’Reilly and Rudy, 2001) place emphasis on different aspects of MTL inputs or computational properties, but each of the dual-process models make relatively consistent predictions with regards to the anatomy of recollection and familiarity.

**1.4.2.2 Other models**

Critics of the dual process theory dispute that recollection and familiarity accurately capture the functional organization of the MTL. In particular, Squire and colleagues suggest that the MTL “operates in a more cooperative fashion than has been envisioned in recent discussions of recollection versus familiarity” (2007).
They account for the data by proposing that the hippocampus processes strong memories, regardless of whether they reflect strong familiarity, recollection, or both (Squire et al., 2007; Wixted and Squire, 2011). They propose that the critical distinction between MTL structures is the type and number of “attributes of experience” they process (Squire et al., 2007). In this view, the hippocampus is involved in combining multiple stimulus attributes, thus supporting both recollection and familiarity when memory requires multi-attribute stimuli (auditory, tactile, temporal). The perirhinal cortex may combine attributes to a lesser extent, e.g., combining visual and spatial attributes.

While a thorough discussion of this possibility is outside the scope of this work, several recent findings argue against this strict “attribute” account of MTL function, if indeed that is what Squire and colleagues are suggesting. One piece of evidence against a purely attribute-centered organization of the MTL comes from crossmodal object recognition. In a series of studies, rats were trained to recognize objects either through touch or visually, and then tested on the objects in the untrained modality (e.g. trained visually and tested tactilely and vice-versa; Reid et al., 2012; Winters and Reid, 2010). While crossed lesions of perirhinal cortex and parahippocampal cortex impaired crossmodal object recognition, hippocampal lesions had no effect. Thus, despite the requirement for poly-modal integration in this task (across two MTL areas), the hippocampus was not critical for tactile-to-visual crossmodal object recognition memory.

Additionally, studies of memory for single-attribute stimuli have found that hippocampal lesions impaired performance. For example, patients with hippocampal lesions were impaired on remembering both synthetic sounds (Squire et al., 2001) and common odors (Levy et al., 2004).
If these stimuli were indeed processed as single-attribute stimuli, both of these results run
counter to the “attribute processing” hypothesis.

1.4.3 Evidence for an anatomical distinction

Evidence to support the anatomical distinction proposed by dual-process theories comes from
several lines of evidence, including neuropsychological studies from functional imaging, patients
with brain lesions, and animal studies.

1.4.3.1 Functional imaging

One technique that has shed light on the neural substrates of recognition memory is functional
neuroimaging. Taken collectively, results from functional imaging support the predictions of the
dual-process model.

A review of the fMRI recognition memory literature by Henson (2005) found results that largely
corroborated the anatomical and functional divisions within the MTL proposed by the dual-
process model. Henson reported a trend for hippocampus and posterior medial temporal cortex
(i.e. parahippocampal cortex) to be activated for encoding and retrieval of source information
and associations between items, while anterior medial temporal cortex (i.e. perirhinal cortex)
tended to be activated in tasks that required item information.

Eichenbaum and colleagues (Eichenbaum et al., 2007) also reviewed the imaging literature on
recognition memory, focusing on the contrast between recollection and familiarity signals in the
MTL. Of the 19 reported contrasts that measured activation related to recollection, 84% reported
hippocampal activation and 58% reported posterior parahippocampal activation
(parahippocampal cortex). In contrast, only 11% of the contrasts related to recollection reported
activation in the anterior parahippocampal gyrus (perirhinal and lateral entorhinal cortices). Of the 15 reported contrasts related to familiarity, 87% reported activation in the anterior parahippocampal gyrus (perirhinal and lateral entorhinal cortices). In contrast, only 27% reported hippocampal or posterior parahippocampal activation.

Critics of the dual-process model interpret these data by proposing that the hippocampus is activated by strong but not weak memories (Squire et al., 2007). They claim that strong memories are associated with increased activity in the hippocampus, irrespective of whether they represent strong familiarity or recollection (Wixted and Squire, 2011).

Other researchers have addressed this critique by using methods to control for the memory strength confound. For example, Montaldi and colleagues (2006) matched memory strength between strong familiarity and recollection by comparing recollected responses with those that were “very familiar.” They found that recollection activated the hippocampus, while increasingly strong familiarity did not progressively activate the hippocampus. Kafkas and Montaldi (2012) confirmed these findings and demonstrated that recollection and familiarity are distinguished by unique visual scanning behavior and pupil dilation. Additionally, recollection activated the parahippocampal cortex and familiarity activated the perirhinal cortex. These results mirror those from other studies that have equated memory strength between strong familiarity and recollection using confidence judgment procedures (sometimes in concert with source judgment; Cohn et al., 2009; Yonelinas et al., 2005; Kafkas and Montaldi, 2014; Johnson et al., 2009; but see: Kirwan et al., 2008; Smith et al., 2011; Wais et al., 2010 for alternate interpretations).
1.4.3.2 Lesion studies

A large body of evidence in support of the dual-process model has come from studies of patients with MTL lesions. Most studies support the dual-process account, with task dissociation methods demonstrating that selective hippocampal damage is associated with impaired recall and spared recognition (Aggleton et al., 2000, 2005; Baddeley et al., 2001; Bastin et al., 2004; Düzel et al., 2001; Holdstock et al., 2005, 2002; Mayes et al., 2002; Turriziani et al., 2004; Vargha-Khadem et al., 1997; Yonelinas et al., 2002) and more comprehensive medial temporal lobe damage is associated with impairment in both recall and recognition (Haist et al., 1992; Hamann and Squire, 1997; Kopelman and Stanhope, 1998).

Process-estimation methods also generally support the dual process model, finding that hippocampal lesions impair recollection while sparing familiarity or impairing it to a lesser degree, while widespread MTL damage impairs both processes (Addante et al., 2012; Aggleton et al., 2005; Blaxton and Theodore, 1997; Knowlton and Squire, 1995; Moscovitch and McAndrews, 2002; Turriziani et al., 2004; Vann et al., 2009; Wais et al., 2006; Yonelinas et al., 2002, 2004).

Patients with hippocampal lesions have a decreased proportion of “remember” responses but an unchanged proportion of “know” responses on remember-know tasks (Aggleton et al., 2000; Blaxton and Theodore, 1997; Knowlton and Squire, 1995; Moscovitch and McAndrews, 2002; Vann et al., 2009; Yonelinas et al., 1998, 2004). The DPSD model interprets this as impaired recollection, which is consistent with the BIC model; the UVSD model interprets these results as a general weakened memory trace.
When tested with the confidence judgment procedure, patients with hippocampal lesions show greater symmetry of the ROC curve than controls (Aggleton et al., 2000; Wais et al., 2006; Yonelinas et al., 2002, 2004). According to the dual-process model, a more symmetrical ROC curve represents impaired recollection (Figure 1-2). In the UVSD model interpretation, a more symmetrical ROC represents weaker memory accompanied by a equal variance of the underlying memory strength distributions for the old and new items (instead of “unequal” variance; Squire et al., 2007). Although the UVSD model can fit the data quite well, the theoretical implications of this result are less clear. Why amnesia should spare memory strength to some degree, yet completely prevent an associated increase in the variance of the two items remains unexplained (Yonelinas and Parks, 2007).

In contrast, patients with widespread MTL damage (including extra-hippocampal areas) show deficits in both recollection and familiarity (Wais et al., 2006; Yonelinas et al., 1998) on the PDP, which is consistent with the dual-process model.

However, other studies of patients with putative selective hippocampal lesions found recall and recognition similarly impaired (Manns and Squire, 1999; Manns et al., 2003a, 2003b; Reed and Squire, 1997; Stark and Squire, 2003). Importantly, the cases inconsistent with the dual-process account have overwhelmingly been composed of patients with anoxic damage. Additionally, many studies on both sides of the debate do not include—let alone quantify—the extent of MTL damage in the patients studied. These factors are important because lesions are often quite variable and the cause of the lesion can have unintended consequences for memory tasks. For example, one review of 67 case reports of hypoxia found that the cortex and basal ganglia are
damaged more frequently than the hippocampus (Caine and Watson, 2000). It also found selective hippocampal damage in only 18% of cases (Caine and Watson, 2000). Additionally, one study found that memory impairment after hypoxia due to cardiac arrest correlates better with reduction in global brain volume than hippocampal volume (Grubb et al., 2000). Thus, care must be taken to characterize the size of the lesion in brain-damaged populations and the extent to which deficits can be explained by global cerebral volume reductions; this type of analysis is rarely performed.

One recent study attempted to address these limitations by quantifying estimates of damage from 19 areas in a cohort of patients with a single etiology (colloid cyst; Vann et al., 2009). They found that the mammillary bodies (part of the extended hippocampal system) were critical for recollection—this is consistent with the dual-process model.

It is important to note that the above studies all compare putative hippocampal damage with widespread MTL damage. Reports of patients with extra-hippocampal damage that spares the hippocampus have been elusive, thus preventing completion of the double dissociation. However, Bowles and colleagues (2007) recently identified a patient with unilateral surgical resection of the amygdala and parts of the entorhinal and perirhinal cortices, leaving the hippocampus intact. This individual displays a behavioral profile opposite to hippocampal-lesioned patients, showing impaired familiarity but preserved recollection on two R-K tasks, the confidence judgment task, and the process dissociation procedure. This behavioral pattern after extra-hippocampal damage is consistent with the dual-process account.
Thus, behavioral testing of recognition memory in patients with MTL lesions has largely supported the dual-process model. In general, studies show that extra-hippocampal damage (in particular perirhinal cortex) is associated with impaired familiarity, while hippocampal damage is associated with recollection impairments. However, without precise data about the location and size of MTL lesions of the participants, the results to date remain inconclusive. In addition, due to a lack of information about the lesion size and location in many patients, it is not clear what the relationship is between the extent of damage to specific neural structures and the deficit in the two processes.

The current study, while conducted in a group with a different etiology of memory impairment, attempts to test the models from this literature by collecting volumetric information about MTL structures in concert with behavioral data.

1.4.3.3 Recollection and familiarity in animals

The other literature that informs our understanding of recognition memory is behavioral neuroscience. For decades, the lesion method has been used to study the neural substrates of recognition memory in animals. The field was inaugurated in the 1970s by attempts to replicate the lesion of patient H.M. in monkeys. Since then, many studies of MTL lesions’ effect on recognition memory in rats and monkeys using the delayed non-match to sample task have shown that the MTL is involved in recognition memory (Beason-Held et al., 1998; Malkova and Mishkin, 2003; Mishkin, 1978; Murray and Mishkin, 1998; Nemanic et al., 2004; Zola et al., 2000; Zola-Morgan et al., 1992).
Although traditional animal recognition memory tasks do not allow for the separation of recollection and familiarity, several strategies have been developed to attempt to dissociate these processes in animals. The first were task dissociation strategies. As when these methods are employed in humans, these strategies rely on attempts to create tasks that isolate one or both of the processes underlying recognition memory. For example, tasks that require only object recognition versus tasks that require linking of object and context, such as an object’s position in a field. These studies in animals have provided more direct evidence than human lesion studies as to the roles of each MTL structure in recognition memory, but direct analogies to the processes of recollection and familiarity are not always straightforward.

With these issues in mind, several groups have attempted to separate and quantify recollection and familiarity in animals by using tasks that closely mirror the behavioral tasks used in humans. ROC curves generated by the confidence judgment procedure display the proportion of hits to false alarms across varying levels of confidence as reported by the subject. If a subject employs a liberal decision criterion, he or she will make a large number of hits (old items identified as old), but will also have many false alarms (new items identified as old), as demonstrated by the upper right of the ROC curve (Figure 1-2). Using a strict decision criterion, a subject will minimize false alarms, but will also decrease the proportion of hits (lower left of the ROC curve).

ROC curves can also be constructed by directly influencing the likelihood that a subject will use a liberal or strict decision criterion. In both humans and animals, increasing reward for correct old items and/or decreasing reward for correct new items will predispose subjects to choose “old” when they are unsure, thus employing a liberal decision criterion. Conversely, similar reward
levels for old and new items biases subjects to use a more conservative decision criterion. Thus, by manipulating response bias through differential reward, ROC curves of animals’ performance can be plotted that mirror the confidence judgment procedure in humans. This procedure, tested in both rats and monkeys, is providing evidence for the existence of two processes in animals.

Using the above procedure to test rats’ memory for different odors, Fortin and colleagues (2004) found that control rats produced ROC curves that were markedly similar to human data (asymmetrical), while hippocampal-lesioned rats’ ROC curves were symmetrical, suggesting that they were only using familiarity. In addition, when the delay was increased to match accuracy between the groups, control rats retained an asymmetrical ROC. This suggests that the symmetry observed in the ROC of hippocampal-lesioned rats is due to a loss of the recollection component, as opposed to a generally weaker memory, as the single-process model would predict. Damage restricted to the medial entorhinal cortex also produced a symmetrical, curvilinear ROC (Sauvage et al., 2010a). These results are especially interesting in light of the medial entorhinal cortex’s putative role in contextual representation.

The recollection component of the ROC was also eliminated when odor recognition memory was assessed under a response deadline. When rats were required to make speeded old/new decisions, the ROC became symmetrical around the diagonal, suggesting that, as in humans, familiarity is available faster than recollection (Sauvage et al., 2010b).

In another experiment, Sauvage and colleagues (2008) developed an associative recognition paradigm where performance depends mainly on recollection in control rats. Stimulus pairs were created from combinations of odors and digging media. Rats were trained to distinguish between
intact and rearranged odor-media pairs, and ROCs were developed by manipulating response bias as described above. This task can be performed through recollection (e.g., associating a lemon smell and wood chips) or though familiarity by unitizing the pairs and employing associative recollection (e.g., lemon-smelling wood chips). In this task, hippocampal-lesioned rats had significantly lower estimates of recollection and higher estimates of familiarity, suggesting they used the latter strategy, while intact rats used the former strategy. In addition, the DPSD model fit the data better than the UVSD model, and model-independent regression analyses supported the predictions of the dual-process and not the single-process model.

The anatomical double dissociation was recently completed with the finding that lesions of the amygdala (which is strongly interconnected with the perirhinal cortex) selectively removes the curvilinear component of the ROC (i.e., familiarity) in the odor recognition memory task (Farovik et al., 2011).

The existence of two processes is also supported by a study of object recognition memory in rhesus monkeys using a response bias manipulation as in the rats (Guderian et al., 2011). Although both the DPSD and the UVSD models fit the data well, 11 out of 12 ROC curves produced by four monkeys tested in this way were better fit by the DPSD model than the UVSD model (one was fit equally well by both models). Additionally, the UVSD model predicts that the ROC curves will always be linear in z-space (zROC), while the DPSD model predicts that the zROCs will be U-shaped. The authors found that the average zROCs were significantly U-shaped and 10 of the 12 individual zROCs were significantly U-shaped or approached
significance. They concluded that the UVSD model could not account for the empirical data, while the DPSD model could.

One final note about process estimates in animals: recollection is often equated with or thought of as a requirement to consciously remember “episodes” from the past. The ability of animals to employ “recollection” has been questioned by arguing that they do not experience episodic memory and the associated “feeling of knowing” (but see Hampton, 2001). However, recollection as conceptualized in the BIC model does not require conscious awareness. Indeed, familiarity, which is also understood to include conscious meta-memory in humans (as demonstrated by the “butcher on the bus” phenomenon), seems to be widely accepted as present in animals (Guderian et al., 2011). In support of this argument, there is evidence that other brain areas (such as the prefrontal cortex) may need to be recruited in order for item and context information to guide conscious behavior in humans (Hannula and Ranganath, 2009; Moscovitch, 2008; Ranganath, 2010b; Ryan et al., 2000).

1.4.3.4 Conclusion

The consensus in the field is that the dual-process models are sufficient to explain object recognition memory (Yonelinas, 2002). To date, results from functional imaging, patients with MTL lesions, and animal lesion studies support the dual-process model. In particular, there is convergent evidence that item familiarity is supported by the perirhinal and lateral entorhinal cortices, “context” is supported by PPHC and medial entorhinal cortex, and recollection is supported by the hippocampus. The current study sought to expand these findings by
investigating the relationship between the volumes of MTL structures and measures of recollection and familiarity in healthy older adults and individuals with aMCI.

1.5 Recognition memory in AD and aMCI

MTL degeneration is a hallmark of AD and aMCI. In particular, entorhinal cortex, perirhinal cortex, parahippocampal cortex, and the hippocampus begin to degenerate very early in aMCI (Devanand et al., 2007). In single-domain aMCI, degeneration is largely confined to the MTL, while patients with multi-domain aMCI have additional widespread loss, showing atrophy similar to that found in AD, although less severe (Seo et al., 2007; Whitwell et al., 2007).

Within the MTL, structural imaging studies also show that the volume differences between MCI patients and controls are larger in the entorhinal cortex than in the hippocampus (Du et al., 2001) and the entorhinal cortex has a larger annual percent loss compared to the hippocampus (Jack et al., 2005). A recent study utilizing high resolution fMRI suggested that lateral entorhinal cortex is the most greatly affected brain area in preclinical AD (Kahn et al., 2004).

There has been less research on perirhinal and PPHC size in MCI and AD, but neuropathological and volumetric studies suggest these areas are also affected, probably early in the disease (Bell-McGinty et al., 2005; Braak and Braak, 1991; Burgmans et al., 2011).

In light of the fact that extra-hippocampal MTL structures (in particular lateral entorhinal and perirhinal cortices) and the hippocampus seem to be affected early in aMCI, the dual-process model would predict that both processes should be affected in aMCI, potentially with familiarity more greatly affected, but certainly more impaired than in healthy aging.
These patterns of degeneration have implications for early detection of aMCI. Various strategies have been attempted to quantitatively and/or qualitatively characterize brain degeneration in early AD, with some success (e.g., Dickerson et al., 2008; Wolk et al., 2013). An alternative strategy is to identify cognitive or behavioral predictors of AD-related brain degeneration, which may be more readily incorporated into neuropsychological testing.

Increasing importance has been placed on identification of cognitive or biological markers that are sensitive and specific to development of aMCI and subsequent AD (Buckner, 2004; Mapstone et al., 2014; Park and Reuter-Lorenz, 2009). Pathological aging, however, has been difficult to separate from cognitive and memory decline associated with normal aging. There is increasing evidence that in addition to differences in the extent of memory impairment between normal and pathological aging, differences may also exist in the type of memory loss exhibited, especially with regard to recognition memory (Wolk et al., 2013). If true, these cognitive changes may offer a useful marker for pathological aging.

1.5.1 Previous research

There is evidence to suggest that recollection and familiarity are differentially affected by normal aging. Most studies report that recollection is reduced in old age, while familiarity is unaffected or less affected than recollection (Davidson and Glisky, 2002; Duarte et al., 2006; Howard et al., 2006; Jacoby, 1999; Jennings and Jacoby, 1997; Light et al., 2004; Parkin and Walter, 1992; Rybash and Hoywer, 1996; Schmitter-Edgecombe, 1999; Toth and Parks, 2006; Yonelinas, 2002).
Preservation of familiarity in old age may be a hallmark of normal aging; in light of the fact that extra-hippocampal MTL structures (in particular lateral entorhinal and perirhinal cortices) and the hippocampus seem to be affected early in aMCI, the dual-process model predicts that both processes should be affected in aMCI, potentially with familiarity more greatly affected, but certainly more impaired than in normal aging. Despite these rather clear predictions suggested by the nature of brain degeneration in AD, studies that have investigated recollection and familiarity in AD and aMCI found inconsistent results (Ally et al., 2009; Anderson et al., 2008; Bennett et al., 2006; Dudas et al., 2005; Hudon et al., 2006; Westerberg et al., 2006; Wolk et al., 2008). In general, all found that recollection is impaired, but deficits in familiarity have been inconsistent.

In AD, there is clear evidence that recollection is impaired (Budson et al., 2000; Christensen et al., 1998; Dalla Barba, 1997; Gallo et al., 2004; Knight, 1998; Koivisto et al., 1998; Smith and Knight, 2002). Some of these studies suggest that familiarity is spared (Dalla Barba, 1997; Lekeu et al., 2003; Rauchs et al., 2007; Tendolkar et al., 1999), but others found impaired familiarity (Algarabel et al., 2009, 2012; Budson et al., 2000; Gallo et al., 2004; Hudon et al., 2009; Smith and Knight, 2002; Westerberg et al., 2006).

Similar to the findings in AD patients, the majority of studies found recollection was impaired in aMCI, while impairment in familiarity was less consistent (Algarabel et al., 2009, 2012; Ally et al., 2009; Anderson et al., 2008; Embree et al., 2012; Hudon et al., 2009; Serra et al., 2010; Westerberg et al., 2006; Wolk et al., 2008, 2013). Many of these studies employed process dissociation methods that allow direct comparisons of process strength and, in these studies, estimations of the magnitude of impairment in familiarity ranged from no impairment to a
greater impairment than in recollection. Of note is the fact that discrepant results were found even in studies using the same type of task.

Table 1-1. Summary of results from studies of recollection and familiarity in aMCI

<table>
<thead>
<tr>
<th>Study</th>
<th>Task</th>
<th>Recollection compared to controls</th>
<th>Familiarity compared to controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algarabel et al. 2009</td>
<td>Task dissociation words</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Algarabel et al. 2012</td>
<td>Task dissociation words</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Westerberg et al. 2006</td>
<td>Task dissociation pictures</td>
<td>↓</td>
<td>=</td>
</tr>
<tr>
<td>Westerberg et al. 2006</td>
<td>Task dissociation pictures</td>
<td>↓</td>
<td>↓</td>
</tr>
<tr>
<td>Wolk et al. 2008</td>
<td>Process dissociation color words</td>
<td>↓</td>
<td>⇩⇩</td>
</tr>
<tr>
<td>Wolk et al. 2008</td>
<td>Process dissociation word-pair</td>
<td>↓</td>
<td>⇩⇩</td>
</tr>
<tr>
<td>Wolk et al. 2013</td>
<td>Process dissociation word-pair</td>
<td>↓</td>
<td>⇩⇩</td>
</tr>
<tr>
<td>Anderson et al. 2008</td>
<td>Process dissociation words</td>
<td>↓</td>
<td>=</td>
</tr>
<tr>
<td>Hudon et al. 2009</td>
<td>Remember-know words</td>
<td>↓</td>
<td>=</td>
</tr>
<tr>
<td>Ally et al. 2009</td>
<td>Confidence judgment words</td>
<td>↓</td>
<td>⇩⇩</td>
</tr>
<tr>
<td>Embree et al. 2012</td>
<td>Confidence judgment pictures</td>
<td>↓</td>
<td>=</td>
</tr>
</tbody>
</table>

⇩ Process is impaired compared to controls; ⇩⇩ Process is more impaired than the other process; = process is unimpaired compared to controls

Each study is summarized briefly, and the results are displayed in Table 1-1. Notably, the vast majority of studies only probed recollection and familiarity in a single task, whereas the strongest evidence for impairment would be replicated over a variety of tasks.
Westerberg and colleagues (2006) used a task-dissociation method to investigate recollection and familiarity in individuals with aMCI. On an item recognition task that required old/new decisions, patients were impaired, suggesting impairment in recollection. On a forced-choice task in which the target was presented in an array with three similar foils, aMCI patients performed similarly to older adults. This task has been held to rely mostly on familiarity (but see Cook et al., 2005; Khoe et al., 2000; Yonelinas, 2002). AD patients were impaired on both tasks. Westerberg and colleagues interpreted their results as demonstrating that aMCI patients had impaired recollection and intact familiarity.

Wolk and colleagues (2008) investigated recognition memory in aMCI using three different tasks: an associative process-dissociation procedure (PDP) task, a featural PDP task, and an item vs. source task-dissociation paradigm. The associative PDP task tested participants on word pairs, one third of which were presented as studied, one third of which were rearranged (previously seen words in novel pairings), and one third of which were completely novel. The featural PDP task tested memory for words studied in green or red font versus novel words. The task dissociation comparison contrasted performance on a standard old/new recognition memory task with performance on a source memory task (which quadrant of the screen the picture was originally presented in).

All three tasks revealed impairment in both familiarity and recollection in aMCI subjects compared to controls, with two of the tasks revealing a greater impairment in familiarity than recollection. Older controls were impaired on recollection alone compared to young adults. In line with the predictions of the dual process model, a composite measure of the three tasks
showed that recollection was correlated with hippocampal volume while familiarity was correlated with extra-hippocampal cortical MTL volume (Wolk et al., 2011). The PDP result was replicated in an additional aMCI population, and they found that familiarity, but not recollection, was correlated with a biomarker of AD measuring cortical thinning (Wolk et al., 2013).

Anderson and colleagues (2008) also used a PDP task to investigate recognition memory in aMCI. Comparing words presented auditorily and visually, they found recollection was reduced in aMCI compared to older controls, but familiarity was unimpaired. This PDP task differed from Wolk and colleagues (2008) in that the learning and testing phases were interspersed by repeating the stimuli at lags of 0, 3, or 12 intervening items, making the study-test delay much shorter. It is possible that this difference caused subjects to use a different strategy that was more reliant on recollection, and thus showed more deficits in that process. Alternatively, familiarity deficits in aMCI may only appear with longer delays.

Hudon and colleagues (2009) used the R-K task to probe memory of 30 single word stimuli presented one at a time. In the testing phase, aMCI and AD participants were shown 30 old and 30 new items and asked to indicate whether the word was old or new, then, if old, to say “remember” or “know” to indicate the quality of their memory. At each trial, participants also had to indicate why they chose that response. Hudon and colleagues found impairment in recollection but intact familiarity for aMCI compared to controls. Participants with AD were impaired in both processes. In this study, “know” responses (as a proxy for familiarity) are reported as a probability that is not corrected for false alarms, instead of as $d'$, which does account for differences in false alarm rates. Considering that aMCI and AD patients in the study
had a higher rate of false alarms than controls, this may have led to overestimating levels of familiarity in these populations and conceal a familiarity deficit in aMCI.

Ally and colleagues (2009; Embree et al., 2012) used standard confidence judgment tasks to investigate memory for words and pictures in aMCI. In Ally and colleagues (2009), during the learning phase the participants were presented with words one at a time and were asked to indicate whether they liked or disliked the item (deep encoding) or to count the number of syllables (shallow encoding). In Embree and colleagues (2012), participants were shown words or photographs of single items, and only the deep encoding condition was employed. During the testing phase in both studies, participants were shown individual items and rated their confidence on a scale from 1-6 that they had seen the item before.

For word stimuli, in both deep and shallow encoding conditions, aMCI subjects showed evidence of impairment in familiarity and recollection. In contrast, picture stimuli demonstrated only a deficit in recollection. The authors attribute this to an intact picture superiority effect in aMCI that can compensate for familiarity deficits.

Serra and colleagues (2010) employed both the standard PDP and R-K tasks with two encoding conditions for each task (words were either read aloud or created through anagrams) to investigate these processes in aMCI. They found discrepant results between the two tasks. In the PDP read-aloud condition, neither process was impaired; in the anagram condition, there was a trend toward statistical significance for worse recollection and familiarity in the aMCI group. In the R-K task, they found a deficit in recollection for both read-aloud and anagrammed words, but no deficit in familiarity. These discrepant results may be due to the exceptionally small number
of stimuli employed (13 in one encoding condition and 12 in the other), or the mismatched number of stimuli per condition (25 old and 30 new), which was not controlled for. Because of the many caveats regarding this study, the results are not displayed in Table 1-1.

In an attempt to resolve the discrepancies above, Algarabel and colleagues (2009) investigated the presence of recollection and familiarity deficits in several populations: young adults, healthy older adults, aMCI patients, non-amnestic MCI patients (naMCI), and AD patients. They used a unique task from Parkin and colleagues (2001) that manipulates perceptual fluency (thought to be a component of the familiarity response). Participants are presented with a series of words composed from two separate groups of letters, and performance is compared with a condition in which all words are composed of letters from both groups. The difference between the two conditions is thought to reflect the use of familiarity arising from the repetition of specific letters.

They found that recollection decreases with age and neurological impairment, but familiarity was only impaired in patients with aMCI and AD, not healthy older adults or naMCI subjects. This suggests that familiarity deficits are specific to the AD process. They also posit that the patients in the Westerburg and Anderson studies were less impaired, and thus less likely to demonstrate familiarity deficits.

In another study, Algarabel and colleagues (2012) examined the performance of young and old controls and AD and aMCI subjects on an associative recognition task to measure recollection and a forced-choice task with foils perceptually similar to those in the associative task to measure familiarity. It is important to note that there is evidence that associative recognition tasks can be performed with familiarity if items are “unitized” into one, e.g. cat-fish remembered
as catfish. In accordance with their previous results, they found that for recollection, older controls were impaired compared to young controls, and the neurologically impaired populations were impaired relative to the older controls. However, they found that familiarity was unimpaired in older controls, but impaired in aMCI and AD subjects. While lending support to the idea that familiarity is impaired in aMCI, the unique nature of the tasks in these studies does not shed light on why discrepant results were seen in studies using the same tasks.

1.5.2 Resolving differences

Several possibilities exist to explain discrepancies in aMCI performance on measures of familiarity. First, heterogeneity among aMCI participants (including severity of impairment, single- or multiple-domain impairments, and underlying pattern of neuropathology) may contribute to variable results. Second, certain tasks may be less suitable for impaired populations (e.g., the R-K and confidence judgment tasks rely on self-report of meta-memory, while the PDP does not), and only a few of the studies investigated multiple tasks in the same participants to compare results across tasks. Third, differences in other task parameters, such as length of time between study and test or type of stimuli used may also affect the results; for example, one group found evidence that familiarity may remain intact for pictures but not words in aMCI (Embree et al., 2012).

1.6 ε4 allele of the apolipoprotein E gene

In addition to cognitive markers of AD, biomarkers—including cerebrospinal fluid markers and PET amyloid imaging—have also been directly correlated with both MTL degeneration and recognition memory. One of the most widely studied genetic biomarkers is the apolipoprotein E
(ApoE) gene, of which the ε4 allele is consistently associated with increased risk of AD (Saunders et al., 1993). Previous research on ApoE-ε4-positive patients with AD and non-demented carriers revealed deficits in both recognition memory (Gilbert and Murphy, 2004) and implicit memory (Negash et al., 2007, 2009). In addition, in patients with AD, the ε4 allele has been associated with increased degeneration in entorhinal cortex (Juottonen et al., 1998), but not the hippocampus (Jack et al., 1998), while ε4 status was associated with hippocampal volume reduction in normal older controls (Lu et al., 2011), though this finding has been inconsistent (Scarmeas and Stern, 2006). Participants in the current study were genotyped for ApoE; if the above associations are correct, the dual-process model predicts an association between ε4 status and familiarity driven by increased entorhinal degeneration.

1.7 The current study

This study sought to quantify the impairment in recollection and familiarity, and determine whether these impairments were correlated with the volumes of the four major MTL structures as predicted by the dual process model.

In this study, we measure recollection and familiarity in aMCI and older adults using three different paradigms: the PDP, R-K, and confidence judgment tasks. All three tasks provide direct measurements of recollection and familiarity, but rest on different assumptions (as discussed above in Separating recollection and familiarity). In addition, R-K, and confidence judgment tasks require introspection about the quality of the participants memories, while the PDP uses a direct (if more limited) measure of recollection. The tasks used similar procedures and stimuli to minimize differences between the tasks. Measures of recollection and familiarity were compared
between tasks to investigate the consistency of results and to lend insight into the nature of the recognition memory deficit in aMCI. This is the first study to employ three different types of tasks in the same aMCI participants. Based on the neurodegeneration associated with aMCI, I predicted that participants with aMCI would have impairments in both processes compared to controls.

In addition, measures of the volumes of four MTL structures: the hippocampus and the entorhinal, perirhinal, and posterior parahippocampal cortices were compared to estimates of recollection and familiarity to explore the relationship between the volumes of these structures and recognition memory performance across controls and participants with aMCI. This is the first study to explore the relationship between these four brain volumes and recognition memory. I predicted that measures of recollection would correlate most strongly with volumes of the hippocampus and PPHC, and that measures of familiarity would correlate most strongly with volumes of the entorhinal and perirhinal cortices.
PARTICIPANTS

2.1 Participant methods

2.1.1 Participant demographics

18 healthy controls (7 male, 11 female) and 15 patients with aMCI (12 male, 3 female) participated in the study. Patients were recruited from neurology and neuropsychology practices in the surrounding community. Diagnosis of aMCI was made by the referring neurologist and/or neuropsychologist, and supplemental neuropsychological testing was completed as part of the study to support the diagnosis. Participants with single- and multiple-domain aMCI were included.

Diagnosis was made according to standard criteria (Petersen, 2004; Petersen et al., 2001; Winblad et al., 2004): subjective memory complaint; objective evidence of memory impairment; intact general cognitive function and activities of daily living; and absence of dementia. There were no strict cut-offs for memory scores, but in general patients performed worse than 1.5 standard deviations below demographically-adjusted norms. Clinical judgment was used in diagnosis, considering a patient’s pre-morbid status and performance on other neuropsychological testing.

Controls were age- and education-matched to the aMCI participants recruited from the community (demographic data are presented in Table 2-1). Controls received the same basic neuropsychological testing as aMCI participants (described below). No controls scored over 1.5 standard deviations below demographically-adjusted norms on neuropsychological testing.
Inclusion criteria for all participants were: aged 50 or over, medically stable, at least 10 years of education, and fluent in English. Exclusion criteria were: history of a neurological disorder (besides aMCI, if applicable); significant psychiatric illness; clinical stroke; traumatic brain injury; alcohol or drug abuse/dependence; contraindications for MRI; or score > 10 on the Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH-D; Williams, 1988). All subjects provided written, informed consent, and were paid for their participation. The study was approved by the Georgetown University Institutional Review Board.

2.1.2 Genotyping

Saliva samples were collected from participants in order to determine participants’ ApoE genotype using the Oragene-DNA collection kit (James et al., 2008). One control participant was not genotyped. Saliva DNA was purified using Oragene purification protocol using prepIT-L2P (OG-L2P-5; DNA Genotek). Total DNA concentration for each sample was determined using a Nanodrop spectrophotometer (Thermo Scientific). ApoE genotype was determined through the Taqman SNP Genotyping Assays and Allelic Discrimination assay (Applied Biosystems; Shen et al., 2009).

2.1.3 Neuropsychological assessment

All participants underwent the following neuropsychological testing: (1) Mini Mental Status Exam (Folstein et al., 1975), (2) Wechsler Memory Scale (WMS) III Logical Memory I and II (Wechsler, 1997), (3) WMS III Word List I and II (Wechsler, 1997), (4) Verbal fluency phonemic (FAS) and semantic (animals; Spreen and Strauss, 1998), (5) Visual Search Task (Goldstein et al., 1973; Lewis and Rennick, 1979), (6) Digit Span Forward (Strub and Black,
2000), (7) Benton Judgment of Line Orientation (Benton, 1994), (8) 60-item version of the Boston Naming Test (Kaplan et al., 1983), (9) Clock-Draw Test (Shulman et al., 1993), (10) Rey Complex Figure (Stern et al., 1999), (11) Trail Making Test A and B (Reitan, 1958), (12) SIGH-D (Williams, 1988). Participants with aMCI (with the assistance of an informant of their choosing) also underwent the Clinical Dementia Rating (CDR) Assessment (Morris, 1993) and Instrumental Activities of Daily Living (IADL) questionnaire (Lawton and Brody, 1969). One aMCI participant was unable to complete the Rey Complex Figure task and another aMCI participant was unable to complete the Benton Line Orientation task. Psychometric data are presented in Table 2-2.

2.1.4 Statistical analyses

Control and aMCI participant demographics and neuropsychological scores were compared using independent samples t-tests. For the neuropsychological assessment, Bonferroni-adjusted p-values \( \frac{0.05}{29} \) set significance at \( p < 0.002 \). Chi-square tests were used to compare expected and obtained ratios.

2.2 Participant results

2.2.1 Participant demographics

Participant demographics are presented in Table 2-1. The mean age of control participants was 69 (s.d. 7) years, and the mean age of participants with aMCI was 70 (s.d. 9) years; ages of the two groups were not significantly different \( [t(31) = 0.36, p = 0.72] \). Years of education were not significantly different: control participants had a mean of 16.8 (s.d. 2.4) years of education, participants with aMCI had an average of 18.3 (s.d. 2.2) years of education \( [t(31) = 1.76, p = \)
0.09]. A chi-square test showed that sex ratios were significantly different between the two groups \( \chi^2(1, N = 34) = 4.48, p = 0.03 \), but there is no evidence that males and females have differential reliance on or abilities in recollection or familiarity.

Being a carrier of the ε4 allele of the ApoE gene is the most significant risk factor for Alzheimer’s disease, after aging (Corder et al., 1993; Verghese et al., 2011). In the general U.S. population, around 27% of the population is positive for at least one ε4 allele; among the AD population, the number is approximately 61% (Raber et al., 2004). In this study, 5 out of 17 control participants (29%) were positive for ε4, while 10 out of 15 participants with aMCI (67%) were ε4 positive, almost exactly the expected percentages. A chi-square test found that the proportions of ε4 positive carriers were significantly different between the two groups \( \chi^2(1, N = 32) = 4.44, p = 0.03 \). These data support the neuropsychological assessment (below) that our two groups accurately reflected control and aMCI populations, respectively.

There were no significant associations between ε4 allele carriage and process estimates or MTL volumes across or within the aMCI and control groups.

Table 2-1. Participant demographics

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>aMCI</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>s.d.</td>
<td>Mean</td>
</tr>
<tr>
<td>Age</td>
<td>69.0</td>
<td>7.7</td>
<td>70.1</td>
</tr>
<tr>
<td>Education</td>
<td>16.8</td>
<td>2.4</td>
<td>18.3</td>
</tr>
<tr>
<td>Sex</td>
<td>7M:11F</td>
<td>12M:3F</td>
<td></td>
</tr>
<tr>
<td>ApoE-ε4 -/+</td>
<td>12:5</td>
<td>5:10</td>
<td></td>
</tr>
</tbody>
</table>

* significantly different from controls, \( p < 0.05 \); n.s. not significantly different
2.2.2 Neuropsychological assessment

Neuropsychological data are presented in Table 2-2. When corrected for multiple comparisons, aMCI participants were significantly impaired on measures of immediate- and long-term episodic (Logical Memory), and verbal (Word List) memory; semantic fluency was also impaired. General cognitive ability as measured by score on the MMSE and executive functioning as measured by the TMT B showed a tendency toward being significantly impaired.

Learning, as measured by learning slope in Logical Memory and Word List, was not significantly impaired in aMCI participants. Visuospatial abilities (Visual Search, Benton, Rey Copy) and spatial memory (Rey Immediate and Delay) were unimpaired. Verbal working memory (Digit Span) was intact as were phonemic fluency and picture naming. Processing speed was also unimpaired (TMT A). Neither group was depressed (score < 10 on HDR), nor were their scores significantly different from each other [t(31) = 1.283, p = 0.21].
Table 2-2. Neuropsychological data

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>aMCI</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>s.d.</td>
<td>Mean</td>
</tr>
<tr>
<td>MMSE</td>
<td>28.2</td>
<td>1.3</td>
<td>26.7</td>
</tr>
<tr>
<td>Instrumental Activities of Daily Living</td>
<td>7.47</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>CDR</td>
<td>0.56</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>WMS Logical Memory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall Total Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning Slope</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recall Total Score II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent Retention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Fluency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonemic - FAS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic - Animals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visuospatial Abilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual Search Correct</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual Search Time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benton Line Orientation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rey Complex Figure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copy Placement and Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Placement and Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Placement and Accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Retention Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Retention Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rey Organization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamilton Depression Rating Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston Naming Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMT A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMT B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Span</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clock Drawing Test</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* significantly different from controls at Bonferroni-corrected level of p < 0.002; † p < 0.005; n.s. not significantly different from controls
EXPERIMENTAL PROCEDURES AND RESULTS

3.1 Overview of experimental sessions

Participants completed the experimental procedures in three sessions of about two hours each. The first and last sessions were completed within one month of each other, but were otherwise unconstrained as to scheduling and order of sessions. Participants completed neuropsychological assessments 1 - 9 above (and aMCI-specific testing, if appropriate) during a single session. In another session, participants completed the three experimental paradigms described below. During the 10 minute delay included in each paradigm, participants completed neuropsychological assessments 10 - 12. In a third session, participants underwent cerebral MR imaging.

Participants were pseudorandomly assigned to complete the three experimental paradigms in a counterbalanced order. Experiments were presented on a computer using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). Two patients and one control are not included in analysis of Experiment 1 (all items endorsed as “remember” or all as “familiar”).

3.2 Experiment 1: Remember-know procedure

3.2.1 R-K methods

3.2.1.1 R-K stimuli

The stimuli were 370 unique color photographs of common objects in natural scenes broken into four lists: two lists of 10 items (for the practice session) and two lists of 175 items (for the experimental session). Assignment of items to each list and the order in which they were
presented was randomized for each participant. Although recognition memory experiments in humans more often use words as stimuli, novel images were used to reduce the potential confound of pre-experimental familiarity.

3.2.1.2 R-K design

Participants first completed a practice session, then immediately completed the experimental section. In both the practice and experimental sessions, the task consisted of five phases: instructions, learning phase, delay, review of instructions, and recognition.

Participants were told they would be seeing a series of pictures that they would be tested on later and they would be asked to make judgments about whether they have seen the item before.

In the learning phase, participants were presented with items to study one at a time (10 for the practice session and 175 for the experimental session). Participants were given five seconds to make a decision of whether they liked or disliked the item (indicated by a key press that advanced to the next item). Participants were told to indicate their preference for the item pictured in the photograph and not the photograph itself, in order to encourage deeper encoding.

The delay was three to five minutes in the practice session and 10 minutes in the experimental session. During the 10 minute delay, participants completed neuropsychological testing. After the delay, the instructions and definitions of terms were reviewed.

“New” was defined as “a picture that has not been presented before.” “Old” was defined as “a picture that has been presented before.” “Remember” was defined as “you can consciously recollect details of the actual occurrence of the picture and can report to the experimenter, if
asked, what you recollected about it.” The term “familiar” was used instead of “know,” as pilot testing indicated participants found it easier to remember the definitions of the terms when “familiar” was used instead of the traditional term “know.” “Familiar” was defined as “you are confident that the picture was presented, but cannot remember anything specific about its presentation.”

The conservative instructions recommended by Rotello and colleagues (2005) and Yonelinas (2001) and a detailed practice session was used to increase the likelihood that “remember” and “familiar” responses accurately reflected recollection and familiarity, respectively. Before testing, participants were asked to explain the instructions in their own words, and were asked to explain their answers on practice trials to ensure their understanding of the task and terms.

During recognition, participants were shown a series of items, half of which were old and half of which were new (20 for the practice session and 350 for the experimental session), and asked to identify the items as “new” or “old.” If they chose “old,” participants were asked to report if they “remembered” the item, or if it was “familiar.” In the practice session, participants were asked to explain their answers to ensure understanding of the task. In the experimental session, participants had 10 seconds to make each decision (indicated by a key press that advanced to the next item), and were not required to explain their answers.

3.2.1.3 R-K analysis

“Remember” responses provide a relatively pure measure of recollection, so recollection is well-represented by the probability (P) that an old item will receive a “remember” response. In contrast, “familiar” responses provide a measure of familiarity only when it occurs without
recollection. The Independence Remember-Know (IRK) procedure was used to calculate familiarity whether or not it occurs with recollection (Yonelinas et al., 1998). In the IRK procedure, the probability of responding “familiar” (henceforth represented with K, after the original “know” response) to an old item is equal to the probability that the old item is familiar (F) and did not get a “remember” response (R), as represented by this formula:

\[ K_{old} = F_{old} (1 - R_{old}). \]

When rearranged, familiarity “hits” can be calculated with:

\[ F_{old} = \frac{K_{old}}{1 - R_{old}} \]

(Yonelinas et al., 1998). Likewise, familiarity “false alarms” can be calculated with the equation

\[ F_{new} = \frac{K_{new}}{1 - R_{new}}. \]

Patients often have different false alarm rates than controls, and this can bias the estimates of familiarity and recollection (Yonelinas et al., 1998). In order to account for response bias in the estimates, recollection was estimated by adjusting the “remember” responses for false alarms. False alarm “remember” responses are subtracted from hit “remember” responses, and then divided by the opportunity the participant has to make a true “remember” response:

\[ R = \frac{R_{old} - R_{new}}{1 - R_{new}}. \]

As in Yonelinas and colleagues (1998), familiarity was measured as the difference between the average familiarity of the old and new items (d’), using:

\[ d' = Z(F_{old}) - Z(F_{new}). \]
Recollection and familiarity scores for the aMCI group were converted into control-referenced z-scores to enable comparison of the degree of impairment in the two processes by subtracting the control mean from each aMCI score and dividing by the control standard deviation.

Control and aMCI participant scores were compared using independent samples t-tests or Mann-Whitney U tests. Paired-samples t-tests were used to compare control-referenced z-scores.

### 3.2.2 R-K results

One control and two aMCI subjects did not complete the task correctly, answering either only “remember” or only “familiar” during the testing phase; they are not included in the analyses.

The proportion of “old” responses to targets (hits) and lures (false alarms) is displayed in Table 3-1. Though patients with aMCI had a lower hit rate and higher false alarm rate than controls, the differences were not significant [hits: t(31) = 1.62, p = 0.12; false alarms: t(31) = 1.44, p = 0.16].

The proportion of “remember” and “familiar” responses to targets and lures are displayed in Table 3-2. From these responses, measures of recollection (probability; P) and familiarity (d’)
were calculated as described above in R-K analysis (Figure 3-2A). Independent t-tests revealed significantly greater recollection for the control participants than the patients with aMCI \([t(28) = 3.10, p = 0.004]\). Familiarity was also greater in controls than patients with aMCI \([t(28) = 2.014, p = 0.05]\).

Next, z-scores were calculated for the aMCI group that were referenced to the control mean and standard deviation. This was done to directly compare the magnitude of impairment between familiarity and recollection by controlling for any difference in difficulty between making judgments based on the two processes (Nunnally and Bernstein, 1994; Yonelinas, 2002). The mean control-referenced z-score was -1.07 (s.d. 0.86) for recollection and -0.64 (s.d. 0.64) for familiarity (Figure 3-2B). A paired-samples t-test revealed that the z-score for recollection trended toward being significantly lower than for familiarity \([t(12) = 2.07, p = 0.06]\), implying that recollection may be more greatly impaired than familiarity in the aMCI group.

Table 3-1. Hit and false alarm rates by group in the remember-know task

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>aMCI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>s.d.</td>
</tr>
<tr>
<td>Hits</td>
<td>0.79</td>
<td>0.13</td>
</tr>
<tr>
<td>False Alarms</td>
<td>0.18</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Table 3-2. Hit and false alarm rates of “remember” and “familiar” responses in the R-K task

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>aMCI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>s.d.</td>
</tr>
<tr>
<td>Remember Hits</td>
<td>0.62</td>
<td>0.18</td>
</tr>
<tr>
<td>Remember FA</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td>Familiar Hits</td>
<td>0.17</td>
<td>0.10</td>
</tr>
<tr>
<td>Familiar FA</td>
<td>0.11</td>
<td>0.10</td>
</tr>
</tbody>
</table>
3.3 Experiment 2: Confidence judgment procedure

3.3.1 Confidence judgment procedure methods

3.3.1.1 Confidence judgment stimuli

The stimuli were 300 unique color photographs of common objects in natural scenes broken into two lists of 150 items, one list to be studied, the other to be used as foils during the recognition task. Distribution of the items between the two lists was randomized for each participant.

3.3.1.2 Confidence judgment design

The task consisted of five phases: instructions, learning phase, delay, review of instructions, and recognition. Participants were first told they will be seeing a series of pictures that they will be tested on later and they will be asked to make judgments about whether they have seen the item before.
In the learning phase, participants were presented with 150 items to study one at a time. Participants were given five seconds to make a decision of whether they liked or disliked the item (indicated by a key press which advanced to the next item). Participants were told to indicate their preference for the item pictured in the photograph and not the photograph itself, in order to encourage deeper encoding.

Then there was a 10 minute delay. During the delay, participants completed neuropsychological testing. After the delay, participants were given instructions on confidence ratings, from 1 (“sure the picture is new”) to 6 (“sure the picture is old”), following the procedure in Yonelinas and Parks (2007). Participants were instructed to spread their responses across the entire scale, and to use every number at least once.

In the recognition phase, participants were shown 300 pictures (half old and half new) one at a time. They were given 10 seconds to report their confidence that an item was presented before on a scale of 1 to 6 using a key press, which advanced to the next item.

### 3.3.1.3 Confidence judgment analysis

In the recognition phase, participants were asked to rate their confidence in having seen the item before on a scale from 1 (certain it was not studied) to 6 (certain it was studied). To create a receiver operating characteristic (ROC) from these confidence ratings, correctly recognized studied items (hits) were plotted against incorrectly recognized lure (unstudied) items (false alarms) for decisions made with increasingly lenient levels of confidence. This was done by plotting hit and false alarm rates as five (x,y) coordinates in a cumulative fashion; i.e., one point
each for hits versus false alarms for confidence level 6, confidence levels 6-5, levels 6-4, levels 6-3, levels 6-2, and levels 6-1.

Responses (1 - 6) to unstudied items were used to calculate false alarm rates and responses to studied items were used to generate hit rates. Thus, the leftmost point on the ROC represents the proportion of “6” (“sure the picture is old”) responses to new items plotted against the proportion of “6” responses to old items. The next point represents the proportion of “6” and “5” responses to new items plotted against the proportion of “6” and “5” responses to old items, and so on. The hit and false alarm rates for all confidence levels (6 - 1) are necessarily 1.0, since they represent the sum of all the probabilities.

The dual-process signal detection model was then fit to the ROC for each participant by minimizing the sum of squared errors using a Microsoft Excel solver routine made available by Dr. Andrew Yonelinas. This model assumes that recollection reflects a threshold retrieval process and familiarity reflects an equal-variance signal detection process.

The hit (H) and false alarm (FA) rates predicted by the model are determined with the following formulas:

\[
H = R + (1 - R) \cdot \Phi \left( \frac{-d'}{2 - c_t} \right)
\]

\[
FA = \Phi \left( \frac{-d'}{2 - c_t} \right)
\]

where \(R\) represents the probability of recollection and \(\Phi\) represents a cumulative normal distribution function at a given level of familiarity (\(d'\)) at each cumulative confidence point (\(c_t\);
i.e. level of response bias). Thus, recollection is represented by the y-intercept and familiarity is represented by the d' of the ROC curve.

The parameter estimates generated from the group averages were compared to the averages of the ROCs obtained by fitting each participant individually; the two results were comparable, which rules out the presence of averaging artifacts. As in Experiment 1, recollection and familiarity scores for the aMCI group were converted into control-referenced z-scores.

The UVSD model was also fit to the ROC for each participant by minimizing the sum of squared errors using a Microsoft Excel solver routine. The UVSD models the distributions of strengths for old and new items as normal distributions controlled by two parameters. The distance between the old and new item distributions is described by d'_F and the standard deviation of the old item distribution is controlled by σ_F.

The hit (H) and false alarm (FA) rates predicted by the model are determined with the following formulas:

\[
H = \Phi\left(c_k, \frac{-d'_F}{2}, \sigma_F\right)
\]

\[
FA = \Phi\left(c_k, \frac{d'_F}{2}, 1\right)
\]

where \(c_k\) is the response criterion for the \(k^{th}\) criterion level and \(\Phi(c, \mu, \sigma)\) is the integral of the cumulative probability density function of a normal distribution with mean \(\mu\) and standard deviation \(\sigma\) over the range from \(-\infty\) to \(c\).
For each participant’s ROC curve, $R^2$ was calculated for the fits of both the DPSD and the UVSD models using the equation:

$$R^2 = 1 - \frac{\text{regression sum of squares}}{\text{total sum of squares}}.$$ 

zROCs were calculated by averaging the z-score at each criterion for every participant. A score of 0 or 1 was corrected using the formulas $\frac{0.5}{n+1}$ or $1 - \frac{0.5}{n+1}$, respectively, before calculating the z-score because $d'$ is undefined when the proportion of responses equals 0 or 1 (Snodgrass and Corwin, 1988). To determine whether the zROC was significantly U-shaped, linear and quadratic equations were fit by minimizing the sum of squared errors using Prism (GraphPad Software).

As in Experiment 1, control and aMCI participant scores were compared using independent samples t-tests or Mann-Whitney U tests. Paired-samples t-tests were used to compare control-referenced z-scores for aMCI participants. Chi-square tests were used to compare expected and achieved ratios.

$F$ tests were used to compare the fits of linear and quadratic equations. When comparing models of varying complexity, the $F$ test investigates whether the relative increase in sum of squares (SS) from the complicated to the simple model is equal to the relative increase in degrees of freedom (DF), using the equation:

$$F = \frac{(SS_1 - SS_2)/(DF_1 - DF_2)}{SS_2/DF_2}.$$

61
Figure 3-3. Schematic of the confidence judgment procedure. In the learning phase, participants were presented items to study one at a time and given five seconds to make a decision of whether they liked or disliked the item. In the recognition phase, participants were given 10 seconds to report their confidence that an item was presented before on a scale from 1 (certain it was not studied) to 6 (certain it was studied).

3.3.2 Confidence judgment procedure results

To calculate hit and false alarm rates for this experiment, confidence ratings of 4, 5, and 6 were collapsed for targets and lures, respectively (Table 3-3). Patients with aMCI had significantly fewer hits than controls \( t(31) = 2.63, p = 0.01 \), as demonstrated by an independent t-test. They also had significantly more false alarms \( t(31) = 2.76, p = 0.01 \), in line with results from other studies that have found that patients with AD show a more liberal response bias than controls (Budson et al., 2006a, 2006b).

The distribution of responses at each confidence level for control and participants with aMCI is displayed in Figure 3-4. Both groups spread their responses across the entire scale of responses, but control and aMCI participants had significantly different distributions of responses \( X^2(5, N = 9858) = 336, p < 0.000 \). In general, controls were more confident of their responses (more likely to respond 1 or 6), while patients with aMCI were less confident (more likely to respond 3 or 4), though post-hoc independent Mann-Whitney U tests revealed significant differences.
between groups only in levels 1 and 3 \((U = 81, Z = 1.95, p = 0.02\) and \(U = 88, Z = 1.7, p = 0.01\), respectively).

From the confidence responses, recollection (probability; \(P\)) and familiarity (\(d'\)) estimates were calculated using the DPSD model as described above in Confidence judgment analysis (Figure 3-5A). Independent t-tests revealed significantly greater recollection for the control participants than the patients with aMCI \([t(31) = 3.49, p = 0.001]\). Familiarity was also greater in controls than patients with aMCI \([t(31) = 2.163, p = 0.04]\).

As in Experiment 1, control-referenced z-scores were calculated for the aMCI process estimates. The mean control-referenced z-score was -1.27 (s.d. 1.09) for recollection and -0.74 (s.d. 0.96) for familiarity (Figure 3-5B). A paired-samples t-test revealed that the z-score for recollection was not significantly different than the z-score for familiarity \([t(14) = 1.42, p = 0.18]\). The results are in line with the R-K task, with both processes impaired, and recollection possibly slightly more impaired, though the difference is not significant in this case.

The UVSD model was also fit to these data. The ROC data and both model predictions are displayed in Figure 3-6AB. Because both models contain two variables, the \(R^2\) value can be used to compare them. Both models fit the data extremely well. Across groups, the mean \(R^2\) for the DPSD was 0.991 (s.d. 0.008), while the \(R^2\) for the UVSD was 0.993 (s.d. 0.009). A paired t-test found that the UVSD fit the data significantly better than the DPSD \([t(32) = 2.21, p = .03]\). However, the effect size (Cohen’s \(d = 0.21\)) is small, and the fit of both models is effectively at ceiling and thus, practically, provide equally good fits.
Next, the ROC data were z-transformed to calculate zROCs (Figure 3-6C). The DSPD predicts that the zROC will be slightly U-shaped, representing the contribution of recollection to high-confidence recognition responses (Yonelinas and Jacoby, 1994). In contrast, the UVSD predicts a linear zROC with a slope less than one, which reflects the ratio of the standard deviations of the distributions of the new and old items.

Linear equations were fit to control \(y = .69x + 1.67, R^2 = 0.990\) and aMCI \(y = 0.83x + 1.22, R^2 = 0.996\) participants zROCs, as were quadratic equations [control zROC \((y = -0.08x^2 + 0.54x + 1.63, R^2 = 0.994)\); aMCI zROC \((y = 0.05x^2 + 0.89x + 1.22, R^2 = 0.998)\)]. The fits of a straight line compared to a quadratic equation for the zROC plots of each group was compared using an F test, and a straight line provided the best fit for each group \([\text{controls}: F(1,2) = 1.43, p = .3547; \text{aMCI}: F(1,2) = 1.58, p = .3355]\). This is not surprising, because in standard item recognition tasks, the U-shape predicted by the DPSD is very slight (Parks and Yonelinas, 2007).

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>aMCI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>s.d.</td>
</tr>
<tr>
<td>Hits</td>
<td>0.82</td>
<td>0.09</td>
</tr>
<tr>
<td>False Alarms</td>
<td>0.17</td>
<td>0.10</td>
</tr>
</tbody>
</table>

** significantly different from controls, \(p < 0.01\)
Figure 3-4. Distribution of responses at each confidence level. Controls and participants with aMCI had significantly different distributions of responses \(X^2 (5, N = 9858) = 336, p < 0.000\); Post-hoc Mann-Whitney U tests: *\( p < 0.05\), **\( p < 0.01\); error bars represent SEM.

![Confidence Response Distribution](image)

Figure 3-5. Process estimates from the confidence judgment procedure. A) Recollection (P) and familiarity (d') estimates by group; B) control-referenced z-scores of recollection and familiarity of the aMCI group. *\( p < 0.05\), ***\( p < 0.001\); error bars represent SEM.

![Process Estimates](image)
Figure 3-6. Fits of the DPSD and UVSD models. A) Fit of the DPSD model to the ROC data from controls and aMCI participants; B) fit of the UVSD model to the ROC data from both groups; C) linear fit of the zROC data from both groups.
3.4 Experiment 3: Process-dissociation procedure

3.4.1 PDP methods

3.4.1.1 PDP stimuli

Ninety-eight color photographs of common objects in natural scenes were divided into three lists of 34, 34, and 30 items (the first and last two items in the two “old” lists serve as buffers to avoid primacy and recency effects and are not included in the analysis). The order and identity of the items in each list was randomized for each participant.

3.4.1.2 PDP design

The design of this study was adapted from Wolk and colleagues (2008). The task consisted of 5 phases: instructions, learning phase, delay, review of instructions, and recognition. Participants were first told they will be seeing a series of pictures that they will be tested on later and they will be asked to make judgments about whether they have seen the item before and which side of the screen the item was presented on.

In the learning phase, participants were presented with 68 items to study one at a time. Items were pseudo-randomly presented either on the left or right side of the screen. Participants were given five seconds to make a decision of whether they liked or disliked the item (indicated by a key press which advanced to the next decision), and then five seconds to indicate whether the item was presented on the left or right side of the screen (indicated by a key press which advanced to the next item). Participants were told to indicate their preference for the item pictured in the photograph and not the photograph itself, in order to encourage deeper encoding.
Then there was a 10 minute delay. During the delay, participants completed neuropsychological testing. After the delay, participants were instructed to call a picture “old” only if it was presented on a certain side (e.g. right). Studied pictures originally presented on the other side (the left side, in this example) and all novel pictures were to be called “new.” Participants were randomly assigned to the left or right-sided condition.

In the recognition phase, participants were shown 90 pictures (1/3 old left, 1/3 old right, and 1/3 new) one at a time in the center of the screen. They were given 10 seconds to make the old/new decision using a key press, which advanced to the next item.

3.4.1.3 PDP analysis

In the PDP, a set of items is created in which recollection (R) opposes familiarity (F). In this case, items originally presented on the side that participants were instructed to endorse as “old” at test represent the “included” items. “Excluded” items were items originally presented on the other side. In this paradigm, “excluded” items will only be incorrectly endorsed as “old” if they were familiar, but the participant was unable to recollect which side they had originally been presented on, using the formulas

\[ p(excluded) = F - FR \text{ or } p(excluded) = (1 - R)F. \]

In the case of “included” items, the probability they will be correctly endorsed as “old” is equal to the probability that the item was either recollected and/or familiar:

\[ p(included) = R + F - RF \text{ or } p(included) = R + (1 - R)F. \]
These equations can then be solved for both R and F:

\[ R = p(\text{included}) - p(\text{excluded}) \]

\[ F = \frac{p(\text{excluded})}{1 - R} \]

This variation of the PDP avoids the confound of potentially different levels of bias that can arise when testing the same participants under two different test conditions.

To account for differences in the rates of false alarms (FA) between the two groups, F is calculated using a measure of discrimination (d') calculated by

\[ Z(\text{familiarity hit rate}) - Z(\text{FA rate}) \]

(Wolk et al., 2008). If a participant had no false alarms, the correction \( \frac{0.5}{n+1} \) was applied because d' is undefined when the proportion of responses equals 0 or 1 (Snodgrass and Corwin, 1988). As in Experiments 1 and 2, recollection and familiarity scores for the aMCI group were converted into control-referenced z-scores.

As in Experiment 1, control and aMCI participant scores were compared using independent samples t-tests or Mann-Whitney U tests. Paired-samples t-tests were used to compare control-referenced z-scores.
Figure 3-7. Schematic of the process dissociation procedure. In the learning phase, participants were given five seconds to make a decision of whether they liked or disliked the item, and then five seconds to indicate whether the item was presented on the left or right side of the screen. In the recognition phase, participants were instructed to call a picture “old” only if it was presented on a certain side (e.g. right). Studied pictures originally presented on the other side (the left side, in this example) and all novel pictures were to be called “new.” They were given 10 seconds to make this decision.

3.4.2 PDP results

The proportion of “old” responses to inclusion, exclusion, and novel stimuli are presented in Table 3-4. Controls, as compared with participants with aMCI, responded “old” to a significantly higher proportion of included items [t(18) = 3.10, p = 0.006], while participants with aMCI had a higher number of false alarms [“old” response to novel item; t(18) = 3.72, p = 0.002] than the controls. Both groups of participants had a similar “exclusion error rate” [“old” response to “excluded” items; t(31) = 2.01, p = 0.84]. From these responses, measures of recollection (probability; P) and familiarity (d’) were calculated for each group as described above in PDP analysis (Figure 3-8A).

Independent t-tests revealed significantly greater recollection for the control participants than the patients with aMCI [t(25) = 2.71, p = 0.01]. Familiarity was also greater in controls than patients with aMCI [t(31) = 4.44, p < 0.000].

As in Experiments 1 and 2, control-referenced z-scores were calculated for the aMCI process estimates (Figure 3-8B). The mean control-referenced z-score was -0.71 (s.d. 0.45) for
recollection and -1.55 (s.d. 1.00) for familiarity. A paired-samples t-test revealed that the control-referenced z-score for familiarity was significantly lower than the z-score for recollection [t(14) = 2.90, p = 0.012], implying that it was more impaired. These results are in the opposite direction to the R-K and confidence judgment tasks. However, aMCI performance on the recollection measure is near floor, thereby possibly concealing the extent of the deficit in that process.

To investigate this possibility, the data were transformed in several ways. First, only the top 50% most accurate experimental stimuli were included in the analysis; second, only the top 50% of participants (by accuracy) in each group were included in the analysis; third, all participants with a score of zero were removed from the analysis; fourth, the data were arcsine transformed to stretch the data at the low end of the distribution. These procedures and analyses can be found in Appendix 1.

All four of these transformations produced similar results, namely that familiarity was more greatly impaired than recollection, though this difference was only significant in the fourth of these analyses. Failure to detect significance could be due to the loss of power to detect a difference caused by the removal of items or participants.

However, regardless of which process is more impaired, it is notable that the results from all three tasks show impairment in both recollection and familiarity. These results are remarkable in light of the variable reports in the literature of a deficit in familiarity. In order to further investigate the relationship between recognition memory processes, volumes of four MTL regions were compared to the process estimates calculated in Experiments 1-3.
Table 3-4. Proportion of “old” responses to inclusion, exclusion, and novel stimuli by group in the process-dissociation procedure

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th></th>
<th>aMCI</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>s.d.</td>
<td>Mean</td>
<td>s.d.</td>
</tr>
<tr>
<td>Old - Inclusion</td>
<td>0.78</td>
<td>0.10</td>
<td>0.57**</td>
<td>0.24</td>
</tr>
<tr>
<td>Old - Exclusion</td>
<td>0.49</td>
<td>0.26</td>
<td>0.51</td>
<td>0.23</td>
</tr>
<tr>
<td>Old – Novel</td>
<td>0.04</td>
<td>0.05</td>
<td>0.17**</td>
<td>0.13</td>
</tr>
</tbody>
</table>

** significantly different from controls, p < 0.01

Figure 3-8. Process estimates from the process dissociation procedure. A) Recollection (P) and familiarity (d’) estimates by group; B) control-referenced z-scores of recollection and familiarity of the aMCI group. **p < 0.01; *** p < 0.001; error bars represent SEM.

3.5 Experiment 4: Relationship between MTL brain volumes and recognition memory

3.5.1 MR imaging methods

All participants underwent MRI safety screening (Shellock). One control participant was unable to complete the MRI due to claustrophobia.

3.5.1.1 Acquisition

Two T1-weighted MPRAGE images were obtained using a 3 Tesla Siemens Trio whole-body MRI system (TE = 2.91 ms, TR = 2300 ms, TI = 900 ms, FOV = 256 mm, 160 slices, slice
resolution 1.2 mm, voxel size 1.2 mm\(^3\)). Images were acquired in the sagittal plane.

### 3.5.1.2 Processing

Cortical reconstruction and volumetric segmentation was performed with the Freesurfer image analysis suite 4.5, which is documented and freely available for download online (http://surfer.nmr.mgh.harvard.edu/). The technical details of these procedures are described in prior publications (Dale and Sereno, 1993; Dale et al., 1999; Fischl and Dale, 2000; Fischl et al., 1999a, 1999b, 2001, 2002, 2004a, 2004b; Han et al., 2006; Jovicich et al., 2006; Segonne et al., 2004).

Briefly, this processing includes motion correction and averaging of the two volumetric T1-weighted images, removal of non-brain tissue using a hybrid watershed/surface deformation procedure (Segonne et al., 2004), automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures (including hippocampus, amygdala, caudate, putamen, ventricles; intensity normalization (Sled et al., 1998), tessellation of the gray matter white matter boundary, automated topology correction (Fischl et al., 2001; Segonne et al., 2007), and surface deformation following intensity gradients to optimally place the gray/white and gray/cerebrospinal fluid borders at the location where the greatest shift in intensity defines the transition to the other tissue class (Dale and Sereno, 1993; Dale et al., 1999; Fischl and Dale, 2000).

Tissue segmentation was checked by investigator LEU. Dura on the inferomedial surface of the brain was routinely miscategorized as gray matter where it abutted the entorhinal cortical surface. It was erased by hand in all subjects and the Freesurfer processing was re-run. This
erasure was necessary to prevent an over-estimation of gray matter in controls, who were more likely to have miscategorized dura than patients with aMCI, due to their larger entorhinal cortices more closely abutting the dura.

### 3.5.1.3 Region of interest tracing

Manual region of interest (ROI) values were determined for four MTL structures: the hippocampus, entorhinal cortex, perirhinal cortex, and parahippocampal cortex. Although good automated ROI volumetry exists for the hippocampus, automated volumetry is not as good at delineating the MTL cortical structures (in Freesurfer, for example, separate automatically determined entorhinal and perirhinal cortex volumes are not available). Therefore, manual ROI tracing was performed on all structures to ensure comparable measures.

One rater (LEU) performed all tracing blinded to subject identity. To determine intra-rater reliability, a random sample of half the cases was retraced and intraclass correlation coefficients [ICC (1,1)] were calculated for each MTL volume (Shrout and Fleiss, 1979). ICC values showed high trace-retrace reliability for hippocampus (0.98), entorhinal cortex (0.94), perirhinal cortex (0.96), and parahippocampal cortex (0.84) volumes.

The hippocampus was traced according to slightly modified instructions from Pruessner and colleagues (2000; Appendix 2). Briefly, the hippocampus was traced in the posterior to anterior direction. In the hippocampal tail, the coronal view is primarily used. The first slice is two slices (~2 mm) posterior to the slice in which the crus of the fornix can be seen fully in profile. The fornix is included in the first slice after the crus of the fornix can be seen in profile. The medial border was the small line of white matter separating the hippocampus and cortex. If this could
not been seen, an arbitrary border was drawn (see Appendix 2). The lateral border was formed by the temporal horn of the lateral ventricle. The superior border was formed by the hippocampal sulcus and the pulvinar.

In the body of the hippocampus, the superior and inferior borders are easily seen in the sagittal plane. The automatic white matter segmentation from Freesurfer was used as the inferior border of the hippocampus. Lateral and medial borders were determined as above. If the medial border was not clear (as above), a straight line with a 45 degree angle was drawn from the most inferior part of the hippocampus medially to the cistern. The hippocampus fissure and uncal cleft were excluded.

At the superior border with the amygdala, the white matter border between hippocampus and amygdala was included in the tracing (sagittal view). In the coronal view, if the white matter border between amygdala and hippocampus was unclear, the line was continued to the ventricle and checked in sagittal/horizontal views. When medial white matter started to form a clear medial boundary and pull in laterally, that was used as the medial boundary for the hippocampus, including superiomedial white matter. In the sagittal view, the alveus served as a landmark for the superior and anterior border of the hippocampus: the hippocampus continued one additional row of pixels anterior to the alveus.

The entorhinal, perirhinal and parahippocampal cortices were traced according to Feczko and colleagues (2009; see Appendix 3). Briefly, these cortices were traced in the anterior to posterior direction. The rostral boundaries of the entorhinal and perirhinal cortices were 3 mm rostral to the slice where the hippocampus can first be seen. The dorsal boundary of the entorhinal cortex
is the sulcus semiannularis. The medial limit was the pial surface and the superiolateral limit is the white matter of the angular bundle. The lateral border was defined at the shoulder of the medial bank of the collateral sulcus without extending the entorhinal cortex boundary into the collateral sulcus. This location served as the medial boundary for the perirhinal cortex. The lateral border of the perirhinal cortex was the innermost point of the collateral sulcus (CS). When there were two CS (interrupted CS) the “main branch” was determined and used.

The caudal limits of the entorhinal and perirhinal cortices were 3 mm caudal to the appearance of the lateral geniculate nucleus. This was also the rostral boundary for the PPHC. The ventrolateral border for the PPHC was localized at the fundus of the collateral sulcus. The dorsomedial boundary of the PPHC marked the point of maximum curvature within the CS and the junction between cortex and the dorsal-medial shoulder of the CS. The first coronal slice caudal to the first slice of the full crus of the fornix was used as a caudal limit for the PPHC (this is also the last hippocampal slice).

3.5.1.4 Analysis

The volumes of structures are reported as structure volume in mm$^3$ divided by total intracranial volume (ICV) multiplied by 1,000,000 for ease of display. Automated values for ICV from Freesurfer were used. Composite measures of recollection and familiarity were calculated by averaging the scores from the three tasks together in order to provide a more stable measure of the two processes.

Differences between groups were investigated using MANOVA or independent-samples t-test. Relationships between structure volume and process estimates were investigated using linear
regression and partial correlation across both control and aMCI participants.

### 3.5.2 MR imaging results

#### 3.5.2.1 Brain volumes

Participants with aMCI had significantly larger intracranial volume (ICV; controls = 1,443,937 mm$^3$, s.d. 168,652; aMCI = 1,604,307 mm$^3$, s.d. 96,848; $t(26) = 3.35, p = 0.003$), most likely due to a greater ratio of males to females in that group. To account for this difference, brain volumes are reported as adjusted for intracranial volume (ICV), calculated by dividing the volume of the structure by the participant’s ICV and multiplying by 1,000,000 for ease of display.

ICV-adjusted volumes of MTL structures for controls and patients with aMCI are displayed in Figure 3-9. A MANOVA of group (control, aMCI) by structure volume (hippocampus, entorhinal cortex, perirhinal cortex, and PPHC) found a trend toward a significant difference between groups [F(4, 27) = 2.3, p = 0.09]. Post-hoc independent t-tests showed that ICV-adjusted entorhinal cortex [$t(26) = 2.34, p = 0.03$] and perirhinal cortex [$t(30) = 2.53, p = 0.02$] volumes were significantly smaller in the aMCI group than in controls. ICV-adjusted hippocampal [$t(26) = 1.32, p = 0.20$] and PPHC [$t(30) = 0.99, p = 0.33$] volumes were not significantly different between controls and aMCI. These results are consistent with previous reports that Alzheimer’s-related pathology begins in extra-hippocampal MTL cortex (Braak and Braak, 1991).
Figure 3-9. ICV-adjusted volumes of MTL structures for controls and patients with aMCI. There was a trend toward a significant difference between groups [F(4, 27) = 2.3, p = 0.09] across hippocampal (HC), entorhinal cortex (ERC), perirhinal cortex (PRC), and posterior parahippocampal cortex (PPHC) volumes. *p < 0.05 post-hoc t-tests; error bars represent SEM.

3.5.2.2 Relationship between brain volume and memory

The relationship between ICV-adjusted volumes of the four MTL structures (hippocampus, entorhinal cortex, perirhinal cortex, and PPHC) with recollection and familiarity across controls and aMCI was examined using several linear regressions. To provide more stability for these measures and minimize the effects of task differences, composite measures of recollection and familiarity were calculated by averaging the estimates produced by the three tasks.

In the first case, the model included the process estimate as the dependent variable and the volume of one of the four MTL structures as the independent variable. The results from each of
these eight regressions are displayed in Table 3-5. Recollection was significantly predicted by both hippocampal and entorhinal cortex volume. Familiarity was significantly predicted by entorhinal cortex volume.

The strong correlations between MTL volumes (Table 3-6) limit the conclusions that can be drawn from the above regressions. I attempted to account for this in two ways: step-wise regression and regression with simultaneous entry of all variables.

To investigate which of the four MTL variables had the most explanatory power with regards to each process estimate, hierarchical regression models were developed in which the MTL variables were added in a step-wise fashion (Table 3-7). With recollection, only hippocampal volume was entered into the model with the highest explanatory power. With familiarity, only entorhinal cortex volume was entered. These results were not dependent on the order in which variables were added to or removed from the model.

To confirm the results of the step-wise regression and attempt to account for shared variance between MTL volumes, with either recollection or familiarity as the dependent variable, a multiple regression analysis was performed in which all MTL volumes were entered simultaneously (Table 3-8). For recollection, the model was significant, and, of the four regions, the hippocampus was the only predictor to trend toward significance ($\beta = 0.4; p = 0.09$). For familiarity, the model was not significant, and no MTL region emerged as a significant predictor, but the structure with the highest $\beta$ value was, as predicted, the entorhinal cortex ($\beta = 0.32; p = 0.19$) while the $\beta$ values for the hippocampus ($\beta = 0.05; p = 0.90$) and PPHC ($\beta = -0.03; p = 0.90$) were much lower.
To explore whether the relationships among the MTL volumes and processes found in the linear regressions were being driven by the significant correlation between recollection and familiarity estimates \( r(31) = 0.60, p < 0.000 \), partial correlations were calculated between each volume of interest and recollection, controlling for familiarity (Table 3-9), and between each volume and familiarity, controlling for recollection (Table 3-10). Only hippocampal volume was significantly correlated with recollection once familiarity scores were controlled for \( r = 0.46, p = 0.01 \).

Though the entorhinal cortex volume trended toward significance \( p = 0.09 \), it was much less highly correlated with recollection estimates \( r = 0.31 \).

For familiarity, no volumes were significantly correlated once recollection scores were controlled for, most likely because of the small sample sizes (e.g., to detect a significant effect with \( r = 0.19 \), a sample size of 349 is required). Despite this, the structures which had the highest \( r \) with familiarity were, as predicted, the entorhinal \( r = 0.19, p = 0.31 \) and perirhinal cortices \( r = 0.11, p = 0.56 \). The volumes of the hippocampus \( r = 0.01, p = 0.97 \) and PPHC \( r = 0.05, p = 0.80 \) were barely correlated with familiarity estimates.
Table 3-5. Single-predictor linear regression models of recollection and familiarity

<table>
<thead>
<tr>
<th>Recollection Composite</th>
<th>Predictor</th>
<th>β</th>
<th>t</th>
<th>$R^2$</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1. HC</td>
<td>0.53</td>
<td>3.41</td>
<td>0.28</td>
<td>11.64</td>
<td>0.002**</td>
<td></td>
</tr>
<tr>
<td>Model 2. ERC</td>
<td>0.46</td>
<td>2.82</td>
<td>0.20</td>
<td>7.93</td>
<td>0.009**</td>
<td></td>
</tr>
<tr>
<td>Model 3. PRC</td>
<td>0.34</td>
<td>1.97</td>
<td>0.12</td>
<td>3.89</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Model 4. PPHC</td>
<td>0.29</td>
<td>1.64</td>
<td>0.08</td>
<td>2.70</td>
<td>0.11</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Familiarity Composite</th>
<th>Predictor</th>
<th>β</th>
<th>t</th>
<th>$R^2$</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1. HC</td>
<td>0.30</td>
<td>1.71</td>
<td>0.09</td>
<td>2.93</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Model 2. ERC</td>
<td>0.39</td>
<td>2.33</td>
<td>0.16</td>
<td>5.45</td>
<td>0.03*</td>
<td></td>
</tr>
<tr>
<td>Model 3. PRC</td>
<td>0.27</td>
<td>1.56</td>
<td>0.07</td>
<td>2.42</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Model 4. PPHC</td>
<td>0.20</td>
<td>1.10</td>
<td>0.04</td>
<td>1.21</td>
<td>0.28</td>
<td></td>
</tr>
</tbody>
</table>

ICV-corrected hippocampal (HC), entorhinal cortex (ERC), perirhinal cortex (PRC), and posterior parahippocampal cortex (PPHC) were included as the only predictors in linear regressions for familiarity and recollection. * p < 0.05; **p < 0.01

Table 3-6. Correlations between MTL volumes

<table>
<thead>
<tr>
<th></th>
<th>ERC</th>
<th>PRC</th>
<th>PPHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>0.66***</td>
<td>0.37*</td>
<td>0.55***</td>
</tr>
<tr>
<td>ERC</td>
<td>-</td>
<td>0.36*</td>
<td>0.41*</td>
</tr>
<tr>
<td>PRC</td>
<td>-</td>
<td>-</td>
<td>0.46**</td>
</tr>
<tr>
<td>PPHC</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

† p < 0.1; * p < 0.05; **p < 0.01; ***p ≤ 0.001; two-tailed (d.f. = 30)
Table 3-7. Hierarchical, step-wise regression models of recollection and familiarity

<table>
<thead>
<tr>
<th>Recollection Composite</th>
<th>Predictor</th>
<th>β</th>
<th>t</th>
<th>$R^2$</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
<td>0.28</td>
<td>11.64</td>
<td>0.002**</td>
</tr>
<tr>
<td>HC</td>
<td>0.53</td>
<td>3.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERC</td>
<td>0.19</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRC</td>
<td>0.17</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPHC</td>
<td>-0.01</td>
<td>-0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Familiarity Composite</th>
<th>Predictor</th>
<th>β</th>
<th>t</th>
<th>$R^2$</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
<td>0.16</td>
<td>5.45</td>
<td>0.03*</td>
</tr>
<tr>
<td>HC</td>
<td>0.07</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERC</td>
<td>0.39</td>
<td>2.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRC</td>
<td>0.15</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPHC</td>
<td>0.05</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ICV-corrected hippocampal (HC), entorhinal cortex (ERC), perirhinal cortex (PRC), and posterior parahippocampal cortex (PPHC) were entered into a hierarchical regression in a step-wise fashion. Model 1 was the most predictive model from the step-wise regression. Shaded rows represent variables not included in the model.

* p < 0.05; ** p < 0.01

Table 3-8. Multiple predictor linear regression models of recollection and familiarity

<table>
<thead>
<tr>
<th>Recollection Composite</th>
<th>Predictor</th>
<th>β</th>
<th>t</th>
<th>$R^2$</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
<td>0.32</td>
<td>3.18</td>
<td>0.03*</td>
</tr>
<tr>
<td>HC</td>
<td>0.40</td>
<td>1.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERC</td>
<td>0.16</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRC</td>
<td>0.17</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPHC</td>
<td>-0.08</td>
<td>-0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Familiarity Composite</th>
<th>Predictor</th>
<th>β</th>
<th>t</th>
<th>$R^2$</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
<td>0.17</td>
<td>1.43</td>
<td>0.25</td>
</tr>
<tr>
<td>HC</td>
<td>0.05</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERC</td>
<td>0.32</td>
<td>1.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRC</td>
<td>0.15</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPHC</td>
<td>-0.03</td>
<td>-0.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ICV-corrected hippocampal (HC), entorhinal cortex (ERC), perirhinal cortex (PRC), and posterior parahippocampal cortex (PPHC) were simultaneously entered into a linear regression model. † p < 0.1; * p < 0.05
Table 3-9. Partial correlation with recollection controlling for familiarity

<table>
<thead>
<tr>
<th></th>
<th>HC</th>
<th>ERC</th>
<th>PRC</th>
<th>PPHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.46</td>
<td>0.31</td>
<td>0.23</td>
<td>0.22</td>
</tr>
<tr>
<td>Significance</td>
<td>0.01*</td>
<td>0.09†</td>
<td>0.20</td>
<td>0.24</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>

† p < 0.1; * p < 0.05; **p < 0.01

Table 3-10. Partial correlation with familiarity controlling for recollection

<table>
<thead>
<tr>
<th></th>
<th>HC</th>
<th>ERC</th>
<th>PRC</th>
<th>PPHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>0.01</td>
<td>0.19</td>
<td>0.11</td>
<td>0.05</td>
</tr>
<tr>
<td>Significance</td>
<td>0.97</td>
<td>0.31</td>
<td>0.56</td>
<td>0.80</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>
Discussion

This is the first study to investigate recollection and familiarity in aMCI using three different tasks: the remember-know task, the confidence judgment procedure, and the process dissociation procedure. As predicted, the data are consistent with impairment in both recollection and familiarity.

In addition, this is the first study to explore the relationship between recognition memory impairment and the volumes of four MTL structures: the hippocampus and the entorhinal, perirhinal, and posterior parahippocampal cortices. The data are consistent with the dual-process model, with the strongest relationships being between recollection and hippocampal volume and familiarity and entorhinal cortex volume.

4.1 Impairment in recollection and familiarity

I predicted that both processes would be impaired in aMCI because the disorder is associated with medial temporal lobe degeneration and AD pathology in the regions that are thought to underlie recognition memory (Whitwell et al., 2007). In particular, lateral entorhinal cortex, perirhinal cortex, and hippocampus are all affected early in the disease (Bell-McGinty et al., 2005; Du et al., 2001; Jack et al., 2005; Kahn et al., 2004). As predicted, the three behavioral tasks revealed similar results, with each task showing clear evidence of impairment in both memory processes, though they differed in the relative magnitude of impairment. In particular, the R-K and confidence judgment tasks found the processes equally impaired while the PDP showed familiarity more impaired than recollection.
4.1.1 Recollection

Impairment in recollection has been found in every study to investigate this process in aMCI, across a variety of different tasks and paradigms (Algarabel et al., 2009, 2012; Ally et al., 2009; Anderson et al., 2008; Embree et al., 2012; Hudon et al., 2009; Serra et al., 2010; Westerberg et al., 2006; Wolk et al., 2008, 2013). Recollection impairment has also been consistently found in studies of patients with mild AD (Algarabel et al., 2009; Budson et al., 2000; Dudas et al., 2005; Gallo et al., 2004; Rauchs et al., 2007; Westerberg et al., 2006). This impairment is consistent with reports of AD pathology in the hippocampus in patients with aMCI and probable AD (Guillozet et al., 2003; Morris and Price, 2001; Morris et al., 2001); given the strong association between recollection memory and the hippocampus, this finding would be expected.

Of note, however, is the fact that aMCI is often diagnosed on the basis of impairment in neuropsychological tests (such as WMS Logical Memory or Word Lists) that test free recall of episodic or verbal memory. Recall performance is thought to depend heavily on recollection, and in task dissociation procedures, is often contrasted with recognition performance (thought to be dependent on both processes) to provide a relatively process-pure measure of recollection. Thus, patients with aMCI may be thought to have deficits in recollection by definition and the finding of a recollection deficit may not be surprising.

4.1.2 Familiarity

Impairment in familiarity has been found much less consistently among studies of aMCI participants (Table 1-1). Across a range of tasks, some studies found no impairment, while others found familiarity more impaired than recollection. Studies of mild AD have been similarly
inconsistent, with some authors reporting a relative sparing of this process and others finding clear evidence of impairment (Bartok et al., 1997; Dalla Barba, 1997; Lekeu et al., 2003; Rauchs et al., 2007; Tendolkar et al., 1999).

In light of recent findings implicating the lateral entorhinal cortex as most greatly affected in early AD (Kahn et al., 2004) and the strong correlation between entorhinal and perirhinal AD pathology (Mitchell et al., 2002), the dual-process model would predict that familiarity should also be impaired in this population, as found in the current study. The inconsistency in previous findings is likely due to a number of reasons, including methodological and task differences and heterogeneity in the patient populations, each explored in further detail below.

4.1.2.1 Methodological differences

While strong conclusions cannot be drawn with regards to the effects of methodological differences due to the many varied paradigms used in the literature, some tendencies emerge. For example, two studies that found preserved familiarity in aMCI, Westerberg and colleagues (2006) and Anderson and colleagues (2008), both used a very short study-test intervals compared to this study and others that found a deficit in familiarity. It is possible that familiarity is more vulnerable to delay in aMCI populations, and only appears on scales of minutes to hours. It is also possible that over shorter delays aMCI subjects use a different mnemonic strategy that is more dependent on recollection, thus masking a familiarity deficit.

Another difference between paradigms is the nature of the stimuli used. In the majority of previously reported studies, the stimuli were single words or word-pairs while the current study used photographs of items in scenes. One study (Embree et al., 2012) directly compared
recollection and familiarity for words and pictures using the confidence judgment procedure. The authors report preserved familiarity for pictures, and attributed this finding to an intact picture superiority effect in aMCI that can compensate for familiarity deficits. They postulated that the comparatively richer conceptual and perceptual information afforded by pictures may drive a stronger sense of familiarity or boost post-retrieval monitoring of familiarity to increase accuracy.

These results are in direct conflict with the results from Wolk and colleagues (2008), who found a familiarity deficit in aMCI using a task dissociation process with pictorial stimuli, and with results from the current study. Differences from the current study are particularly interesting, as both employed the confidence judgment procedure, similar numbers of stimuli, and equivalent study-test interval lengths. It is possible that the stimuli employed in the current study (photographs of items in scenes), despite being even more conceptually and perceptually rich than the stimuli used in Embree and colleagues (2012; single-item color photos), were more similar to each other, thus making the task more taxing to familiarity and more sensitive to familiarity deficits.

4.1.2.2 Task differences

While the wide variety of tasks available to investigate and quantify recollection and familiarity has been a boon to the field, allowing converging evidence from paradigms with varying assumptions to support and reinforce each other, they have also led to difficulties in reconciling discrepancies between results when they occur. One strength of the current study is the use of three different measurement methods (the remember-know task, the confidence judgment
procedure, and the process dissociation procedure) in the same group of participants; it is the first study in aMCI to do so. The only other study to employ more than one type of task in the same set of aMCI participants is Wolk and colleagues (2008), who used two PDP tasks and a task dissociation procedure. Assessing performance on more than one task allows one to rule out differences in subject characteristics as the root of performance differences. As mentioned earlier, it is notable that the results from the three tasks in this study were in concert because, while all three tasks provide direct measurements of recollection and familiarity, they rest on different assumptions.

In particular, two of the tasks require introspection about the quality of the participants memories, while the third uses a direct (if more limited) measure of recollection. One possible concern is the suitability of self-report tasks for impaired populations, but the fact that the results from two different types of self-report (confidence and remember-know) were in line with a task that uses an objective measure of recollection suggests that participants were able to accurately judge the quality of their memories. It is also notable that in the confidence judgment task, participants with aMCI and controls had similar levels of false alarms at confidence level “6” (controls = 0.06 ± 0.07; aMCI = 0.07 ± 0.07; t(31) = 0.35, p = 0.73), suggesting that aMCI subjects were able to accurately identify when they were confident they had not seen an item before. Future studies might employ response bias manipulations (as in animal behavior studies) in this population to compare estimates generated from self-report with those that don’t require meta-memory.
In addition, during training before the R-K task, all participants were able to define, in their own words, the difference between “remember” and “familiar” memories and were able to provide justification for choosing “remember,” suggesting that at the very least, these categories were understood by participants.

Two studies have employed the confidence judgment procedure in aMCI: Embree and colleagues (2012), discussed above, which used photographic stimuli and found impairment only in recollection, and Ally and colleagues (2009), from the same group, which found deficits in both familiarity and recollection using a similar number of stimuli as in the current study (though the items were words). The current study also found impairment in both processes. Differences between the three studies may be a result of the nature of the stimuli employed.

The only other study which successfully investigated recognition memory in aMCI using the remember-know task found that recollection was impaired but familiarity was not (Hudon et al., 2009). However, in this study, “know” responses were not corrected for false alarms and aMCI participants had a higher rate of false alarms than controls; this could overestimate levels of familiarity in this population. Interestingly, patients with AD tested on the same task did show impairment in familiarity. Thus, one explanation for the differential findings might be a generally lower impairment in their patients with aMCI compared to the current study, although the reported MMSE is higher in that study than the current one, which makes this explanation less likely. One last difference between the tasks is the number of stimuli used. The current study employed 175 target items while the Hudon study only used 30 words. Thus, in the current
study, the study-test interval was significantly longer and there was greater potential for interference between stimuli—factors that may tax familiarity more than recollection.

The PDP task, while offering an objective measure of recollection, has a caveat in that only specific types of recollection are measured. In this task, only recollection about the side of the screen the target was originally presented on was measured. Not assessing so-called “non-criterial recollection,” or recollection that occurs but does not help inform performance on the PDP (such as remembering that you coughed when the item was first presented), may artificially deflate recollection scores and inflate familiarity scores. Notably, estimates of recollection for both control and aMCI participants were lower on the PDP task than in the R-K and confidence judgment procedure, while familiarity scores were higher. However, there is no reason to believe that non-criterial recollection is more common in one group than the other, so the magnitude of the impairment should not be affected.

Finally, these tasks also share the assumption that recollection and familiarity operate independently of each other. Other groups have proposed that the two processes may be redundant (all recollected items are also familiar) or exclusive (items are either recollected or familiar; Mayes et al., 2007); if either of these possibilities is true, estimates of familiarity in particular would be affected. Redundancy would increase the difference in familiarity scores between groups while exclusivity would decrease it. While previous investigations of the independence assumption have largely supported it (Yonelinas, 2002), future studies may investigate methods that do not require the independence assumption (e.g., task dissociation methods as in Algarabel et al., 2009, 2012; Wolk et al., 2008).
4.1.2.3 **Heterogeneity of population**

Another possible explanation for differences seen between studies is heterogeneity of the aMCI population. Participants may differ on the severity of their memory impairment and the extent of AD neuropathology. Because clinical neuropsychological measures are rather gross measures of memory performance, and because there are no standard diagnostic tests for aMCI-related memory impairment, it is difficult to compare the degree of memory impairment in the aMCI groups included in different studies. The most common psychometric reported, the MMSE, is a very crude measure of general cognitive functioning, which doesn’t have the sensitivity or specificity needed to ascertain where in the disease progression participants may fall. These considerations have implications for the distribution of neuropathology, and thus the nature of the impairment.

Another consideration is the number of participants included in each study who will eventually convert to AD. Individuals with aMCI, though more likely to convert to AD than other groups, occasionally revert back to normal or their condition remains stable. The makeup of each particular group will influence memory performance.

Finally, studies also differ on whether both single- and multiple-domain aMCI participants are included and the exact cognitive domains involved also vary. In this study, both variants of aMCI were included, and scores of executive functioning (TMT B), semantic fluency (animals), and general cognitive ability (MMSE) were lower than controls, though the mean score was not always significantly different from the control group. Involvement of these domains may impair recollection or familiarity to different extents, and thus lead to differences between studies.
Future studies employing larger numbers of subjects may investigate the influence of various cognitive impairments on recollection and familiarity.

Familiarity, conceptualized as a more automatic process, has generally been found to be less affected by encoding manipulations that recollection (Yonelinas, 2002), and thus might be postulated to be less affected by such confluent deficits than recollection. However, there is evidence that other brain areas besides the MTL are involved in recognition memory decisions. For example, the lateral prefrontal cortex is involved in familiarity decisions (Yonelinas et al., 2005), and damage to this structure also impairs familiarity (Aly et al., 2011). The lateral prefrontal cortex is especially implicated in the monitoring or decision processes that attend an “old”/“new” decision based on memory strength. Recollection, on the other hand, activates anterior medial prefrontal cortex (Yonelinas et al., 2005), and is impaired in patients with medial and polar prefrontal lesions (Wheeler and Stuss, 2003). The prefrontal cortex may be involved in controlled or strategic search of memory during recollection. Thus, either process might be affected by concomitant deficits in other aspects of memory, depending on the distribution of AD pathology.

One limitation of the current study due to its small sample size is that only the relationships between the two processes and MTL brain areas were investigated. Future studies could investigate the role that degeneration in other brain areas, such as the prefrontal cortex, may play in producing aMCI memory deficits through voxel-based morphometric studies of the relationship between recollection and familiarity and structure volume over the entire brain. In
addition, functional imaging during encoding and/or retrieval processes may shed light on these considerations and the brain areas implicated in these deficits.

4.1.3 Relative magnitude of impairment

One difference in results between the tasks in the current study that merits discussion is the magnitude of the impairment of the two processes. The R-K and confidence judgment procedure suggested that the processes are equally impaired (with recollection potentially more impaired), while the PDP showed that familiarity was more impaired. This last finding comes with the caveat that aMCI recollection estimates were close to floor, and thus could be obscuring a greater deficit. Of note, however, is the fact that four separate transformations of the data to mitigate the floor effect failed to reverse the direction of the impairment (Appendix 1).

A finding of greater impairment in familiarity could be related to assumptions and methodology inherent to the PDP; for example, non-criterial recollection included in the familiarity measurement may artificially inflate the familiarity deficit, though it is important to note that the magnitude of the recollection deficit found in the other two tasks (z-scores of -1.07 and -1.27) is still smaller than the deficit found in familiarity with the PDP (z-score of -1.55), so this is unlikely to be the only explanation.

Another possibility is that aMCI participant’s ability to introspect about recollection and familiarity differs. Recollection may be more difficult to self-monitor or aMCI participants may be less confident about the strength of their memories, and thus recollection seems more impaired in the two tasks that rely on this method of reporting.
Regardless of which process is more greatly impaired, it is important to note that both recollection and familiarity were impaired, a finding that is discussed further in the next sections.

### 4.1.4 Implications for treatment

Early treatment of mild AD is associated with better outcomes than late treatment (Winblad et al., 2006); thus, identification and treatment of aMCI may be the most effective strategy for preventing or delaying AD onset (Mariani et al., 2007). Given the possibility that patients with mild cognitive impairment may benefit as much or more from AD treatments than those already diagnosed with dementia, many interventions are currently being explored, both pharmacological and cognitive in nature.

Pharmacological agents that are already approved for use in AD, such as acetylcholinesterase inhibitors, have shown limited or no efficacy in delaying progression to AD in aMCI (Simon et al., 2012). One study of donepezil and vitamin E found that treatment delayed conversion to AD at one but not three years (Petersen et al., 2005). However, currently there is no evidence that pharmacological therapies exist for aMCI that provide long-term benefits (Diniz et al., 2009; Raschetti et al., 2007; Simon et al., 2012).

In recent years, cognitive interventions have garnered much attention for normal elderly, particularly memory and attention training (Mariani et al., 2007). These therapies have included restorative training strategies, which train the affected domain directly; cognitive stimulation, which aims to increase functioning through non-specific interventions; and compensatory strategies, which teach participants how to compensate for lost or impaired functions (Jean et al., 2010; Simon et al., 2012). All three approaches have been successful at improving and
maintaining functioning in normal adults (Kueider et al., 2012; Reijnders et al., 2013; Schneider and Yvon, 2013). Studies have even showed maintenance of training effects at five (Willis et al., 2006) and ten years (Rebok et al., 2014).

Several groups have reported successful cognitive interventions in aMCI, though these programs have been hampered by methodological limitations such as small sample sizes (Belleville et al., 2011; Buschert et al., 2010; Jean et al., 2010; Li et al., 2011; Stott and Spector, 2011). There is now ample evidence that aMCI patients are able to learn both new information and compensatory strategies, though the heterogeneity and often non-specific nature of the interventions makes mapping the boundaries of effective interventions difficult (Huckans et al., 2013; Simon et al., 2012). One additional advantage of cognitive interventions is that they can be utilized without worry about medication interactions or side effects that can be a concern with use of cholinesterase inhibitor (Jean et al., 2010).

Future studies may investigate the utility of interventions focused on training in recognition memory, perhaps with a particular focus on restorative or compensatory strategies in the familiarity domain. For example, a restorative strategy was successfully employed by Jennings and colleagues (2005) to train older adults to improve their recollection abilities. In addition, prophylactic studies that attempt to slow decline in specific cognitive domains may be investigated.

4.2 Anatomy of recognition memory

Recent elaborations of the dual-process model have included a possible anatomical segregation between the perirhinal and lateral entorhinal cortices that support familiarity and item
information, and the hippocampus, PPHC, and medial entorhinal cortex, which are thought to underlie recollection and contextual information. This is the first study to explore the relationship between recognition memory and measures of the volumes of four MTL structures: the hippocampus and the entorhinal, perirhinal, and posterior parahippocampal cortices.

In line with the predictions of the dual-process model, the current study found that familiarity was most greatly associated with entorhinal cortex volume and recollection was most associated with hippocampal volume.

Although the entorhinal cortex was also associated with recollection when each MTL area was separately regressed with recollection, the strong correlations between MTL volumes limit the conclusions that can be drawn from these independent regressions. Thus, two additional regressions were performed to account for these relationships (step-wise regression and regression with simultaneous entry of all variables). These additional regressions gave similar results: that the hippocampus was most associated with recollection and the entorhinal cortex was most associated with familiarity.

An additional consideration is the correlation between recollection and familiarity—when familiarity was added as a nuisance covariate, the relationship between entorhinal cortex volume and recollection also disappeared. Thus, either of the inter-correlations (between volumes or between processes) may have been driving that association. Future studies will benefit from larger sample sizes, so that these relationships may be teased apart despite the inter-correlations between the different measures and volumes.
Even if the correlation between the entorhinal cortex and recollection is not spurious, it might be expected that they would be associated, given that the entorhinal cortex provides the major source of inputs to the hippocampus. This fact has sometimes been cited as a challenge to the independence assumption for recollection and familiarity. However, the independence of the two processes can still be accounted for by recognizing that the familiarity of an object is not the only information provided to the hippocampus by the perirhinal and entorhinal cortices. In addition to the familiarity of an object, there is also evidence that the perirhinal cortex has a role in high-level object perception (Bartko et al., 2007; Bussey et al., 2002; Murray et al., 2007), and provides the hippocampus information about the properties of the object. The dual-process model postulates that even if the perirhinal and entorhinal cortices are unable to signal the familiarity of the object, enough information about the properties of the object is sent to the hippocampus to enable pattern completion and subsequent recollection. However, if the extra-hippocampal cortex is too damaged, the hippocampus will essentially be functionally disconnected, and recollection will be prevented. This failure of recollection is not due to a loss of familiarity, per se, but might depend on the perceptual functions of MTL cortex.

Another interesting consideration is the relative degeneration in medial and lateral entorhinal cortex, which are thought to be differentially associated with recollection and familiarity, respectively. Previous research suggests that the greatest area of neuropathology in early AD is the lateral entorhinal cortex (Kahn et al., 2004), thus degeneration in the lateral entorhinal cortex may drive the correlation with familiarity. Unfortunately, lateral and medial entorhinal cortex are very difficult to differentiate through anatomical tracing, and thus the differences between these areas cannot easily be determined using volumetric MRI as in this study. Future studies with
larger sample sizes might perform a relatively crude 50:50 split to divide the medial and lateral entorhinal cortex and investigate differences with respect to recollection and familiarity.

It might seem expected that the aMCI population, to some extent selected for impairment in recollection, would have large hippocampal volume loss, and thus show a significant correlation between the two variables. However, it doesn’t necessarily follow that this deficit must be due to hippocampal damage. As discussed above, recollection impairment can also follow from frontal damage or from deficits in encoding or post-retrieval monitoring, and not necessarily from a deficit in memory, per se. Moreover, the relationship is significant even in the control group alone ($R^2 = 0.24$, $\beta = 0.49$, $F(1, 15) = 4.76$, $p = 0.05$), suggesting it is not an artifact driven by the aMCI group.

Somewhat surprisingly, the perirhinal cortex was not especially strongly associated with familiarity (though it was one of the only variables that was correlated once recollection was factored out). This may be due to the relative difficulty in accurately determining the boundaries of entorhinal and perirhinal cortex using anatomical landmarks. These structures are usually defined cytoarchitectonically, and, although the ROI tracings in the current study follow established protocols and were consistently drawn, there is no way to independently confirm the accuracy of the boundaries for each subject.

The results from this experiment are consistent with the majority of results in the amnestic literature that have found that hippocampal lesions impair recollection, while more widespread lesions impair both processes (Holdstock et al., 2002; Yonelinas et al., 2002). The double dissociation in the amnestic literature was recently completed with the report of a patient with
isolated left surgical resection of perirhinal cortex and parts of the entorhinal cortex and amygdala (Bowles et al., 2007). This patient has impaired familiarity with preserved recollection. However, it is just a single case, and, as MTL injury sparing the hippocampus is a relative rarity, similar cases have yet to be reported.

Also consistent with the results from this experiment is a study by Yonelinas (2007), who investigated the two processes in normal aging, finding that recollection was associated with hippocampal volume and familiarity was associated with entorhinal volume. These two processes were measured using a task-dissociation procedure, with recall performance thought to depend mostly on recollection and item recognition thought to depend on familiarity. They supported this dissociation with structural equation modeling.

In the only other study to investigate these relationships, Wolk and colleagues (2011) found a stronger relationship between recollection and hippocampal volume and between familiarity and extrahippocampal MTL structure volume (perirhinal, entorhinal, and posterior parahippocampal cortex combined). These process estimates were composite measures created from the results from two variants of the PDP.

These results underscore the potential utility of the familiarity measure for aMCI and AD, as it may differentiate between healthy and pathological aging, an important goal in this field. While recollection is impaired in healthy aging, familiarity is relatively spared (Davidson and Glisky, 2002; Yonelinas, 2002). Familiarity impairment may also be relatively specific to aMCI, and not be simply a general marker of pathological aging. For example, other disorders of aging, such as vascular dementia (Tierney et al., 2001) and hippocampal sclerosis are more likely to have
impairment in recollection with spared familiarity. In addition, the single study to date that has investigated these processes in non-amnestic MCI reported that this population also had impaired recollection with spared familiarity (Algarabel et al., 2009).

Supporting this theory, a recent study measured estimates of familiarity and recollection in young adults, controls, and participants with aMCI, and correlated the strength of each process with a structural imaging biomarker of AD (Wolk et al., 2013). Called the “cortical signature of AD,” this biomarker measures cortical thickness in several areas of the brain affected in AD, and has been previously demonstrated to correlate with symptom severity in preclinical AD (Dickerson et al., 2008). This study found that familiarity was significantly correlated with this biomarker, while recollection was not. Familiarity was more highly correlated with the AD signature than any other standard neuropsychological measure, suggesting that it may be sensitive to AD-specific brain changes early in the disease.

As greater emphasis is placed on early diagnosis, calls for better characterization and definition of the nature of the memory impairment in aMCI have increased (Dubois and Albert, 2004; Dubois et al., 2007). Free recall, which is often used to diagnose aMCI, is more likely to depend on recollection, which could be affected by hippocampal dysfunction or frontal-subcortical dysfunction, and is unlikely to distinguish between various etiologies of amnesia. In addition, current clinical tests that rely primarily on familiarity often fail to distinguish deficits in aMCI because of ceiling effects (Wolk et al., 2008). Tests such as those used here, on the other hand, may more accurately distinguish between types of memory disorder. Definitively linking a familiarity deficit in aMCI with prodromal AD will require following aMCI participants
longitudinally to determine who converts to AD. Future studies following subjects longitudinally may also investigate whether volume-function relationships differ with different stages of the disease.

4.3 Effect of ApoE-ε4 status

Although our aMCI population was enriched in ε4 allele carriers, there were no significant associations between ε4 allele carriage and process estimates or MTL volumes across or within the aMCI and control groups. Based on previous research which found an association between ε4 allele presence and increased degeneration in entorhinal cortex (Juottonen et al., 1998), I predicted an association between ε4 status and familiarity estimates driven by larger entorhinal degeneration. It is likely that the current study was underpowered to detect such a relationship, if it exists. Only 5 control participants were ε4 positive and only 5 aMCI participants were ε4 negative, severely limiting the chance of finding significant associations with this measure.

4.4 UVSD model

Finally, the results from this study must be discussed in the context of the model used to acquire them. The DPSD model is just one of many models of recognition memory. Its utility as a model rests greatly on the large body of data collected in many different populations and its ability to provide quantitative estimates of recollection and familiarity, but that doesn’t mean it is the best model of these processes or the only one that can explain the data.

In this study, the UVSD was able to provide a slightly better fit to the ROC data than the DPSD (though practically, the fits were equal; $R^2 = .993$ versus .991). Although the UVSD as modeled cannot produce estimates of recollection and familiarity, many theorists now view the UVSD as
a dual-process model in which familiarity and recollection are summed prior to a recognition memory decision. There are a few models based on the UVSD “summed” model that provide quantitative process estimates. Future studies may investigate the process estimates derived by these models and compare whether they produce process estimates similar to the DPSD, and whether these estimates are likewise affected in memory-compromised populations. Other models of recognition memory may yet provide more utility in identifying, assessing, or predicting memory impairment in aMCI.

As is, this study suggests that the DPSD can provide useful information about how memory is affected in pathological aging, and the relationship between these changes and accompanying neurodegeneration.

### 4.5 Conclusion

In conclusion, using three separate recognition memory tasks, the present study found that both recollection and familiarity are impaired in aMCI, and that measures of these processes are correlated with the volumes of MTL structures as predicted by the dual-process model.

This is the first study to investigate recollection and familiarity in aMCI using three different tasks in the same set of participants. Assessing performance on more than one task allows one to rule out differences in subject characteristics as the root of performance differences. It is notable that the results from the three tasks in this study were in concert because, while all three tasks provide direct measurements of recollection and familiarity, they rest on different assumptions. Previous discrepancies in the literature with regard to aMCI deficits in these processes are likely a result of a combination of factors, including methodological, task, and participant differences.
This was also the first study to explore the relationship between recognition memory and measures of the volumes of four MTL structures. The results from this study are consistent with the dual-process model and suggest that measures of familiarity, in particular, may have utility as a proxy for AD neuropathology, but the predictive value of this measure requires longitudinal study.
APPENDIX 1. PDP floor effect transformations

For the first manipulation, accuracy for each experimental stimulus was calculated across all subjects. Accuracy was defined as “hits” and “correct rejections” across the “inclusion” and “new” categories of stimuli. The “exclusion” category was not included in the measure of accuracy, as participants will only correctly accept an item in this category on the basis of familiarity in the absence of recollection, but will correctly reject an item on the basis of recollection or forgetting. Thus, inclusion of this category might bias the results in favor of familiarity.

A median split was performed on the item accuracy data, and recollection and familiarity estimates were calculated for the top 50% of stimuli in each category (Figure A1-1A). Accuracy for these stimuli was greater than 64%. This manipulation increased estimates of recollection in the aMCI group from 0.1 (s.d. 0.12) to 0.15 (s.d. 0.15), but did not completely remove the floor effect. The familiarity estimate increased from 1.36 (s.d. 0.83) to 1.69 (s.d. 0.96) in the aMCI group. Control-referenced z-scores were then calculated for the aMCI process estimates. The mean z-score was -1.00 (s.d. 0.52) for recollection and -1.44 (s.d. 1.49) for familiarity (Figure A1-1B). A paired t-test found no significant difference [t(14) = 1.05, p = 0.31], but the direction of the effect was unchanged.

For the second manipulation, only the top 50% of participants in each group as measured by overall accuracy in the “inclusion” and “new” categories of stimuli (see above for rationale) were included in the analysis (Figure A1-1C). The mean recollection estimate in the aMCI group
increased to 0.15 (s.d. 0.26), and the mean familiarity estimate increased to 1.64 (s.d. 0.51).

Control-referenced z-scores calculated for the aMCI group (Figure A1-1D) estimated recollection at -0.91 (s.d. 0.54) and familiarity at -1.43 (s.d. 0.76). A paired t-test found no significant difference \([t(7) = 1.29, p = 0.24]\), but the direction of the effect is unchanged.

In the third manipulation, all participants with a recollection score of 0 were removed from the analysis (4 control and 5 aMCI participants; Figure A1-1E). The mean recollection estimate in the aMCI group increased to 0.15 (s.d. 0.12), and the mean familiarity estimate increased to 1.28 (s.d. 0.82). Control-referenced z-scores calculated for the aMCI group (Figure A1-1F) estimated recollection at -0.91 (s.d. 0.49) and familiarity at -1.71 (s.d. 1.49). A paired t-test found no significant difference \([t(10) = 1.55, p = 0.16]\), but the direction of the effect is unchanged. However, this manipulation may bias the relationship between recollection and familiarity scores by selectively removing participants with low recollection.

In the fourth manipulation, the data were arcsine transformed using the equation

\[
2 \times \left(\sqrt{\frac{\text{score}}{\text{maximum score}}}\right)
\]

to stretch the low end of the distribution that is skewed due to the floor effect (Rietveld and Hout, 1993). Recollection and familiarity are displayed in Figure A1-1G, but due to the arcsine transformation, are no longer presented as \(P\) and \(d'\). Control-referenced z-scores calculated for the aMCI group (Figure A1-1H) estimated recollection at -0.63 (s.d. 0.59) and familiarity at -1.50 (s.d. 1.12). A paired t-test found that familiarity was significantly more impaired \([t(14) = 2.62, p = 0.02]\).
Figure A1-1. PDP data transformations. A) Top 50% most accurate items; B) control-referenced z-scores for aMCI group; C) top 50% of participants in each group; D) control-referenced z-scores for aMCI group; E) all participants with a recollection score of 0 removed; F) control-referenced z-scores for aMCI group; G) arcsine transformation of all data; H) control-referenced z-scores for aMCI group. * p < 0.05; error bars represent SEM.
APPENDIX 2. Hippocampal ROI tracing

Hippocampal tracing followed Pruessner (2000), with slight modifications.

The first slice (Figure A2-1A) is one slice posterior to the slice after the crus of the fornix can be seen fully in profile and tracing continued anteriorly (Figure A2-1B). The superior white matter is included starting from the first slice after the crus of the fornix can no longer be seen fully in profile (Figure A2-1C).

For the hippocampal body, the superior and inferior border was perceptible in the sagittal orientation (Figure A2-1D). The fimbria was included and marks the superomedial level of the hippocampal body. One level of gray matter pixels superior to the fimbria is also included. The hippocampal fissure was excluded.

When the subiculum was detached from the cortex and a line of white matter was visible between the hippocampus and parahippocampal cortex or entorhinal cortex, this was the inferomedial border of the hippocampal body (Figure A2-2A). When the border was not clear, a straight line at a 45 degree angle was drawn from the most inferior part of the hippocampal body medially to the cistern (Figure A2-2B).

The lateral border was identified by the inferior horn of the lateral ventricle or the caudally adjacent white matter. If the ventricle itself was not visible, one row of gray matter pixels was excluded from the CA region at this point, assuming it represented the lateral ventricle. In most case the superior excess of the quadrigeminal cistern could be identified as the superomedial border of the hippocampal body.
The alveus serves as a landmark for the superior and anterior border of the hippocampus: The hippocampus continues one additional row of pixels anterior to the alveus in the sagittal view (Figure A2-3). The medial and inferior borders of the hippocampal head are identified as with hippocampal body (use coronal and horizontal views; Figure A2-4A,B).

In the superior border with the amygdala, the white matter border between the hippocampus and the amygdala was included in the tracing (using the sagittal view). If it was unclear, the line was continued to the ventricle and checked in the horizontal view.

When the medial white matter started to form a clear medial boundary and pull in laterally, it marked the medial boundary for the hippocampal head, including superiomedial white matter.
Figure A2-1. **Tracing the hippocampal head.** A) Tracing begins one slice posterior to the slice after the crus of the fornix can be seen fully in profile; B) crus of fornix fully in profile; C) white matter of fornix is included starting from the first slice after the crus of the fornix can no longer be seen fully in profile; D) superior and inferior border perceptible in the sagittal orientation.
Figure A2-2. Tracing the hippocampal body. A) A line of white matter visible between the hippocampus and parahippocampal cortex or entorhinal cortex marks the inferomedial border of the hippocampal body (red); B) when the border was not clear, a straight line at a 45 degree angle was drawn from the most inferior part of the hippocampal body medially to the cistern.

Figure A2-3. Superior and anterior borders of the hippocampus. The hippocampus continues one additional row of pixels anterior to the alveus in the sagittal view.
Figure A2-4. Tracing the hippocampal head. The medial and inferior borders of the hippocampal head are identified as with hippocampal body; A) coronal view; B) horizontal view.

Figure A2-5. Border between amygdala and hippocampus. A) Sagittal view; B) coronal view.
Figure A2-6. Medial boundary for the hippocampal head marked by white matter.
APPENDIX 3. Entorhinal, perirhinal, and posterior parahippocampal cortex ROI tracing

The rostral boundary of the entorhinal cortex and perirhinal cortex was 3 mm rostral to rostral boundary of the hippocampus. Amygdala, white matter of the posterior parahippocampal cortex (PPHC), gyrus ambiens, and sulcus semianularis could be seen in this slice. In the rostral part, the superior limit of the entorhinal cortex was the shoulder of the endorhinal sulcus (point of greatest curvature; Figure A3-1). Tracing continued posteriorly.

Where the uncal sulcus was visible, the superior limit of the entorhinal cortex was the most medial point of the inferior part of the uncal sulcus unless the junction between subiculum and cortex can be seen then that marked was the superior limit (Figure A3-2).

The inferior border of the entorhinal cortex was the middle point of the medial bank of the collateral sulcus. The medial border was the pial surface. The lateral border of the entorhinal cortex was defined at the shoulder of the medial bank of the collateral sulcus without extending the entorhinal cortex boundary into the collateral sulcus. This location served as the medial boundary for the perirhinal cortex. The superiolateral borders of the entorhinal and perirhinal cortices were the white matter surface identified by Freesurfer (Figures A2-1 and A2-2; yellow line).

The ventrolateral border for the perirhinal cortex and PPHC was defined by the lateral branch of the collateral sulcus and localized at the fundus of the collateral sulcus. When there were two collateral sulci (“interrupted collateral sulcus”) the main branch of the collateral sulcus was used. This was identified using the partially inflated pial surface in TKSurfer. If the branch split and
reconverged, the most lateral sulcus was used. If the branch was interrupted and continued in a completely separate branch, the tracing switched to the other sulcus at the slice where the second sulcus appears. (Figure A3-3)

The caudal limit of the entorhinal and perirhinal cortices was 3 mm caudal to the appearance of the lateral geniculate nucleus (Figure A3-4). This was also the rostral boundary for the PPHC. The first coronal slice caudal to the first slice of the full crus of the fornix was used as a caudal limit for the PPHC. The caudal boundary of the PPHC was the same as the hippocampus (Appendix 2)—one slice posterior to the slice after the crus of the fornix can be seen fully in profile.
Figure A3-1. The rostral boundary of the entorhinal and perirhinal cortices. Entorhinal cortex is traced in red; perirhinal cortex is traced in purple; red arrow marks the shoulder of the endorhinal sulcus, the superior limit of entorhinal cortex; yellow lines outline the white matter as identified by Freesurfer.

Figure A3-2. Entorhinal and perirhinal cortex, continued. Entorhinal cortex is traced in red; perirhinal cortex is traced in purple; hippocampus is traced in blue; red arrow marks uncal sulcus; blue arrow marks the shoulder of the medial bank of the collateral sulcus, the border between entorhinal and perirhinal cortex.
Figure A3-3. **Collateral sulcus variations.** Main sulcus marked by red line. A) Uninterrupted; B) partially interrupted; C) interrupted twice; D) split at rostral end—main (lateral) sulcus marked by white asterisk.

![Image of collateral sulcus variations](image)

Figure A3-4. **Appearance of the lateral geniculate nucleus.** LGN marked 3 mm rostral to the caudal limit of entorhinal (purple) and perirhinal (green) cortex and the rostral limit of the PPHC.

![Image of lateral geniculate nucleus](image)
WORKS CITED


McClelland, J.L., McNaughton, B.L., and O’Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. 102, 419–457.


