DISTORTIONS OF SHORT-TERM MEMORY:
FALSE MEMORY, SEMANTIC INTERFERENCE, AND FAMILIARITY

by

Alexandra S. Atkins

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in The University of Michigan
2009

Doctoral Committee:

    Professor Patricia A. Reuter-Lorenz, Chair
    Professor John Jonides
    Professor Douglas Noll
    Assistant Professor Cindy Lustig

Henry Gustav Molaison
February 26, 1926 – December 2, 2008

AKNOWLEDGEMENTS

I would like to thank my advisor, Patricia A. Reuter-Lorenz, for her inspiration and support throughout my graduate career. This work could not have been accomplished without her superb encouragement and collaboration. I am thankful to Prof. Reuter-Lorenz both professionally and personally, and look forward to many years of continued collaboration and friendship.

Next, I must thank Marc G. Berman, John Jonides, and Rick Lewis for their collaborative contributions to the research presented in Chapter 4 of the present dissertation. In particular, I am thankful to Marc for his insights, both scientific and personal. Ours is friendship I will value for years to come.

I would also like to extend thanks to all members of my dissertation committee, Patricia Reuter-Lorenz, John Jonides, Cindy Lustig, and Douglas Noll, for their impeccable service and support throughout the dissertation process.

On a personal note, I would like to thank my closest friends Kristin Scherrer, Jenn Mann, Debs Cane, and Dave Parker. Each of you has changed my life. Each of you has made me laugh, brought me joy, and been a source of encouragement, support and love without which I could not have accomplished this degree. I would also like to thank my parents, Beth and Harry, as well as my brother Daniel and nephew Owen. Thank you for tolerating my absences, welcoming me home with open arms, and for your continued love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACC	accuracy
ANOVA	analysis of variance
AS	articulatory supression
BA	Broadman's area
CAT	categorized
CR	correct rejection
DRM	Deese-Rodeger-McDermott paradigm
FA	false alarm
fMRI	functional magnetic resonance imaging
ms	milliseconds
MTL	medial temporal lobe
NEG	negative
NRN	non-recent negative
NRPOS	non-recent positive
PI	proactive interference
POS	positive
PPC	posterior parietal cortex
PSC	percent signal change
RN	recent negative
ROI	region of interest
RP	recent probes
RP	recent positive
RPOS	recent positive
RT	response time
SE	standard error
SI	semantic interference
ST	short-term
VLPFC	ventro-lateral prefrontal cortex

ABSTRACT

Decades of research have demonstrated that episodic memory is vulnerable to significant semantic distortion (Gallo, 2006). Recent findings suggest that short-term memory is susceptible to similar distortions of meaning. The present investigations explore the cognitive and neural mechanisms of memory distortions that emerge within a few seconds of encoding. Findings demonstrate false recall and recognition of unstudied lure items only 3-4 seconds following encoding of a short, 4-item memory set, and show that correct rejections of lures are associated with considerable semantic interference (SI). An fMRI investigation of these effects suggests a distinction between the left mid-ventrolateral prefrontal cortex (L VLPFC), which shows increased activity changes associated with increased SI, and the right posterior parietal cortex (R PPC) which shows increased activity associated with declines in SI. An investigation of interactions between SI and proactive interference (PI) in short-term memory shows that vulnerability to PI is mediated by the semantic relationship between recently studied items and current memoranda. Taken together, findings are consistent with unitary, activation-based models of memory (Nairne, 2002), and reveal the considerable vulnerability of verbatim memory processes, even over very short retention intervals.

Chapter I

INTRODUCTION

Opening remarks

Human memory is arguably our most precious faculty. Our memories allow us to consider and respond to stimuli no longer present in our sensory environment, learn from past experience, and form meaningful relationships. Generally, we trust that our recollections provide an accurate record of our experiences. Decades of research have shown, however, that memory is not akin to an internal video of the past.   Early work by Bartlett (1932) spoke directly to the vulnerability of memory, demonstrating that recall for unfamiliar narratives can be influenced by the cultural context of readers. Readers make both errors of commission, in which they insert unstudied information into memorized passages, and errors of omission in which they fail to reproduce memoranda that are inconsistent with their expectations.   Bartlett was followed thirty years later by Underwood (1965) who demonstrated the associative nature of memory by showing that subjects were more likely to incorrectly recognize words that were related in meaning to those presented in a continuous stream of memoranda. Underwood (1965) interpreted this finding as evidence for an associative network through which unstudied items were internally generated via spreading semantic activation, an interpretation consistent with later demonstrations of semantic priming (Meyer & Schvaneveldt, 1971), and with models of associative memory (Anderson & Bower, 1973).

1

Modern investigations of memory errors can generally be separated into one of two broad categories: those that investigate semantic distortions of episodic and autobiographical memory, and those that examine interference-related distortions of short-term memory. One reason for this separation is that dominant multi-store models of memory have assumed a relatively strict division between the psychological and neural mechanisms subserving short- and long-term memory (see Jonides, Lewis, Nee, Lustig, et al., 2008). Semantic, or meaning-based processing, has been generally considered the provenance of episodic memory, and errors related to meaning have been attributed to this system regardless of the time-frame in which they occur. Perceptually-based visual and phonological codes, on the other hand, have been considered the hallmarks of short-term memory (Baddeley, 1966, 1972, 1986).

Distortions of episodic memory

In the last two decades, there has been remarkable growth in the study of episodic memory distortions, and the term *false memory* has been widely used to describe instances in which our memories fail to represent events as they truly occurred. Striking examples of this vulnerability come from findings that indicate even memories for highly salient life-altering events are often distorted. For example, most people will report remembering the precise circumstances under which they learned of catastrophic historical events such as the assassination of U.S. President Kennedy or the terrorist attacks of September 11, 2001. However, research on the topic has shown even memories such as these, which are often reported in rich sensory detail and with high

confidence (Brown & Kulik, 1977), are often distorted and subject to change over time (Neisser, 1982, 1986; Neisser & Harsch, 1990; Lee & Brown, 2003; Talarico & Rubin, 2003; Greenberg, 2004; Wolters & Goudsmit, 2005; Conway, Skitka, Hemmerich, & Kershaw, 2009; Hirst, Phelps, Buckner, Budson, et al., 2009). Furthermore, such memories sometimes include factual inaccuracies, which provide indisputable evidence of distortion. A famous example comes from recounts of September 11th, 2001 made by then U.S. President Bush, and countless others, who confidently reported having seen the first airplane strike the World Trade Center on that day, an event that was not televised (Greenberg, 2004).

Other studies of autobiographical memory have shown that recollections of childhood experiences are surprisingly vulnerable to intrusions, particularly when misleading information is introduced by a credible source (Loftus & Pickrell, 1995; Loftus, 1997, 2003). The study of autobiographical memory is challenging, however, as it is difficult to maintain experimental control over stimuli and retention intervals. Except in rare cases, it is also difficult to prove an autobiographic memory is spurious. For this reason, many investigators have turned to techniques that allow them to examine false recall and recognition of memoranda presented in controlled laboratory settings.

In the Deese-Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995), participants study lists of words, all related in meaning to a common theme word, which is unstudied. At test, subjects are then asked to either recall or recognize these memoranda. For example, their first study of this kind, Roediger and McDermott (Roediger & McDermott, 1995) presented subjects with lists of related memoranda at a rate of one word every 1-2 seconds. A subject might study *nose, breathe,*

sniff, aroma, hear, see, nostril, whiff, scent, reek, stench, fragrance, perfume, salts, and *rose.* At the end of study list, some participants freely recalled memoranda, while others engaged in a distracter task for 2 minutes prior to studying the next list. Several minutes after study of the 16th and final list, all subjects were then given a recognition task which asked them to identify studied items from a list of probes that included studied words, unstudied words that were the themes associated with studied items (*smell*, in our example), and unstudied words that were not semantically associated with memoranda. Recall results showed that veridical recall was best for the first and last few presented items, consistent with a standard serial position curve (Waugh & Norman, 1965; Atkinson & Shiffrin, 1968). False intrusions of theme words were made at a rate that rivaled correct recall for items presented toward the middle of the list. Rates of false recognition were also high. Regardless of whether they had engaged in free recall, participants falsely recognized theme items at a rate that far exceeded their false alarms to unassociated probes. In some cases, the false alarm rate to lures even approached the hit rate for items that were actually studied, suggesting participants could not distinguish between the two.

Subsequent research has revealed the robustness of the false memory phenomenon across multiple modalities and stimulus sets. False memory effects occur with both visually and aurally presented word lists, as well as picture stimuli (Roediger, McDermott, & Robinson, 1998; Gallo, McDermott, Percer, & Roediger, 2001; Roediger, McDermott, Pisoni, & Gallo, 2004; Gallo, 2006 for review). Furthermore, though verbatim recognition decreases with increasing retention intervals, rates of false recognition either remain stable or increase over time (Gallo, 2006), a finding that has

been interpreted as evidence for a dissociation between familiarity-based and item-specific processes in recognition memory (see Yonelinas & Levy, 2002; Diana, Reder, Arndt, & Park, 2006).

Distortions of short-term memory

The term *false memory* has been used almost exclusively in reference to meaning-based distortions of episodic memory, as described above. There is, however, ample evidence for intrusions and distortions in short-term memory. For example, when asked to immediately recall visually presented letters, subjects are vulnerable to acoustic intrusions, such as replacing a studied *B* with a similar sounding item, such as a *C* or *P* (Conrad, 1964). When subvocal rehearsal is prevented by the introduction of articulatory suppression, these phonological errors are eliminated (Murray, 1968; Peterson & Johnson, 1971). Visual similarity effects have also been demonstrated in both recognition and recall paradigms, with subjects erroneously reporting having studied words that were visually similar to memoranda (Arons & Mason, 1999; Logie, Della Sala, Wynn, & Baddeley, 2000).

Short-term memory is also vulnerable to proactive interference. For instance, in short-term item recognition, subjects are both slower and less accurate in rejecting negative probes that appeared as memoranda on a previous trial (Monsell, 1978), a finding that has been taken as evidence that a controlled interference resolution process may be recruited to support of accurate retrieval from short-term memory in such cases (see Jonides & Nee, 2006 for review).

Phonological and visual interference effects have been taken as evidence that phonological and visual representations are the primary codes utilized in short-term memory, a view that is most strongly evident in Baddley's model of working memory (Baddeley & Hitch, 1974; Baddeley, 1986). This model assumes that maintenance of information in short-term memory is carried out by storage buffers specialized for phonological and visual-spatial processing, the contents of which can be accessed and manipulated by a central executive.

Semantic influences on short-term remembering have been more difficult to quantify. Although early investigations showed some evidence of semantic interference (e.g. Baddeley, 1966a, 1966b; Dale & Gregory, 1966; Kintsch & Buschke, 1969; Shulman, 1970; Baddeley & Levy, 1971; Shulman, 1972) these findings were largely dismissed in light of strong evidence supporting a dominant role for phonological codes (Baddeley, 1986). As such, well-documented effects such as semantic proactive interference in immediate recall (Wickens, 1973) and semantic influences on memory span (Baddeley, 1966b; see also Hulme, Maughan, & Brown, 1991) have not been considered distortions of short-term memory, despite the time frame in which they occur.

More recent work has questioned the notion that semantic coding is unique to long-term memory. While some have argued for distinct semantic codes in short-term memory (Hamilton & Martin, 2005; Martin, Shelton, & Yafee, 1994), others have questioned the veracity of multi-store models altogether (e.g. Cowan, 2001; McElree, 2001; Oberauer, 2002; Ranganath & Blumenfeld, 2005; Jonides, et al., 2008).

Unitary vs. multi-store models of memory

While multi-store models argue for a psychological and neural dissociation between short- and long-term memory, unitary models generally argue the contents of short-term memory may be best described as the activated contents of long-term memory (Cowan, 2001; McElree, 2001; Oberauer, 2002; Jonides, et al., 2008). Early support for the neural dissociation of short- and long-term memory came from observed double dissociations between patients with medial temporal lobe lesions, who appeared to exhibit selective deficits in episodic memory (Scoville & Milner, 1957), and those with damage to left perisylvian regions, who demonstrate deficits in phonological processing and short-term retention (Warrington & Shallice, 1969). However, recent accounts have questioned the selectivity of these deficits (see Ranganath & Blumenfeld, 2005). Furthermore, neuroimaging investigations have revealed considerable overlap between the neural regions supporting short- and long-term remembering in neurologically intact subjects (e.g., Cabeza, Dolcos, Graham, & Nyberg, 2002; Ranganath, Johnson, & D'Esposito, 2003; Karlsgodt, Shirinyan, van Ep, Cohen, & Cannon, 2005).

Although such findings propose a challenge to dual-storage accounts, overlap between the neural correlates of short- and long-term memory are consistent with unitary accounts which assume only a single set of memory representations. Although unitary models differ with respect to the number of representations that may be active simultaneously (see Jonides, et al., 2008), most assume that the representations themselves are multidimensional. For example, according Nairne's (2002) feature-based

model, memory representations consist of a constellation of activated cues corresponding to task-relevant aspects of memoranda, such their phonological, lexical and semantic features. These activated cues are then used to reconstruct the memoranda at retrieval.

Overview of the present dissertation

Over the past 40 years the assumption that semantic representations are unique to episodic memory may have diverted interest away examining the extent to which we are vulnerable to rapid distortions of meaning. Given the potential influence of such distortions on immediate behavior across a variety of contexts, this is a topic of considerable practical importance.

The present research directly addresses this topic by examining the rapid acquisition of false memories and semantic interference. In Chapter II, I introduce a short-term variation on the DRM paradigm (Deese, 1959; Roediger & McDermott, 1995), and demonstrate that false recognition and recall of unstudied lure items can occur only seconds following encoding. I also find evidence that short-term remembering is vulnerable to considerable semantic interference effects, wherein the correct rejection lures is slowed relative to correct rejection of unrelated items.

In Chapter III, I go on to examine the neural mechanisms underlying these semantic distortions. I suggest a dissociation between the left ventrolateral prefrontal cortex and right intra-parietal sulcus regions, with the former showing increased activation associated with increases in semantic interference, and the latter demonstrating increased activation associated with reduced vulnerability to false recognition.

Chapter IV presents an investigation of the relationship between proactive and semantic interference in short-term memory. Here, I demonstrate that the semantic context of probes presented in an item recognition task can change the degree to which we are vulnerable to proactive interference from previously studied information. In my concluding remarks (Chapter V), I summarize findings, offer some additional theoretical considerations, and explore future directions.

Taken together, this body of work provides important insights into the mechanisms underlying rapid distortions of memory. By incorporating approaches from multiple domains of memory research, this work will contribute to the development of more accurate models of cognitive control and action in the real world. Furthermore, the research serves as a step towards revealing the circumstances that may increase our vulnerability to memory errors (e.g., distraction, aging, pathology), and may therefore contribute to the development of interventions to prevent such errors.

REFERENCES

Anderson, J. R., & Bower, G. H. (1973). *Human Associative memory*. Washington, D.C.: Winston and Sons.

Arons, S. E., & Mason, A. (1999). Effects of visual similarity on serial report and item recognition. Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 52(1), 217-240.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A proposed system and its control process. In K. S. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation* (Vol. 2, pp. 89-195). Oxford: Academic Press.

Baddeley, A. D. (1966a). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18A*, 362-365.

Baddeley, A. D. (1966b). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18*, 302-309.

Baddeley, A. D. (1986). *Working Memory*. Clarendon: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., & Levy, B. A. (1971). Semantic coding and short-term memory. *Journal of Experimental Psychology, 89*, 132–136.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge Univ. Press.

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage, 16*(2), 317-330.

Conrad, R. (1964). Acoustic Confusions in Immediate Memory. *British Journal of Psychology, 55*(1), 75-84.

Conway, A. R. A., Skitka, L. J., Hemmerich, J. A., & Kershaw, T. C. (2009). Flashbulb Memory for 11 September 2001. *Applied Cognitive Psychology, 23*(5), 605-623.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci, 24*(1), 87-114; discussion 114-185.

Dale, H. C. A., & Gregory, M. (1966). Evidence of semantic coding in short-term memory. *Psychonomic Science*, 75-76.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17-22.

Diana, R. A., Reder, L. M., Arndt, J., & Park, H. (2006). Models of recognition: a review of arguments in favor of a dual-process account. *Psychon Bull Rev, 13*(1), 1-21.

Gallo, D. A. (2006). *Associative Illusions of Memory: False Memory Research in DRM and Related Tasks*. New York: Psychology Press.

Gallo, D. A., McDermott, K. B., Percer, J. M., & Roediger, H. L. (2001). Modality effects in false recall and false recognition. *Journal of experimental psychology. Learning, memory, and cognition, 27*(2), 339.

Greenberg, D. L. (2004). President Bush's false 'flashbulb' memory of 9/11/01. *Applied Cognitive Psychology, 18*(3), 363-370.

Hirst, W., Phelps, E. A., Buckner, R. L., Budson, A. E., Cuc, A., Gabrieli, J. D. E., et al. (2009). Long-term memory for the terrorist attack of September 11: Flashbulb memories, event Memories, and the factors that influence their retention. *Journal of Experimental Psychology-General, 138*(2), 161-176.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*(6), 685-701.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annu Rev Psychol, 59*, 193-224.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139*(1), 181-193.

Karlsgodt, K. H., Shirinyan, D., van Ep, T. G. M., Cohen, M. S., & Cannon, T. D. (2005). Hippocampal activations during encoding and retrieval in a verbal working memory paradigm. *Neuroimage, 25, 25*(4), 1224-1231.

Kintsch, W., & Buschke, H. (1969). Homophones and synonyms in short-term memory. *Journal of Experimental Psychology, 80*(3P1), 403-407.

Lee, P. J., & Brown, N. R. (2003). Delay related changes in personal memories for September 11, 2001. *Applied Cognitive Psychology, 17*(9), 1007-1015.

Loftus, E. F. (1997). Creating false memories. *Scientific American, 277*(3), 70-75.

Loftus, E. F. (2003). Make-believe memories. *The American Psychologist, 58*(11), 867.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25*(12), 720-725.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 53*(3), 626-646.

McElree, B. (2001). Working memory and focal attention. *J Exp Psychol Learn Mem Cogn, 27*(3), 817-835.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.

Monsell, S. (1978). Recency, Immediate Recognition Memory, and Reaction-Time. *Cognitive psychology, 10*(4), 465-501.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679-684.

Neisser, U. (1982). Snapshots or Benchmarks? In U. Neisser (Ed.), *Memory Observed: Remembering in Natural Contexts*. San Francisco: Freeman.

Neisser, U. (1986). Remembering Pearl Harbor - reply. *Cognition, 23*(3), 285-286.

Neisser, U., & Harsch, N. (1990, Feb 02-03). *Phantom flashbulbs: False recollections of hearing the news about Challanger.* Paper presented at the Conference on Affect and Accuracy in Recall, Atlanta, GA.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology-Learning Memory and Cognition, 28*(3), 411-421.

Peterson, L. R., & Johnson, S. T. (1971). Some effects of minimizing articulation on short-term retention. *Journal of Verbal Learning & Verbal Behavior, 10*(4), 346-354.

Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences, 9*(8), 374-380.

Ranganath, C., Johnson, M. K., & D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia, 41*(3), 378-389.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lIsts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803.

Roediger, H. L., McDermott, K. B., Pisoni, D. B., & Gallo, D. A. (2004). Illusory recollection of voices. *Memory, 12*(5), 586-602.

Roediger, H. L., McDermott, K. B., & Robinson, K. J. (1998). The role of associative processes in creating false memories. In M. Conway, S. Gathercole & C. Cornoldi (Eds.), *Theories of Memory II* (pp. 187-245). Hove, Sussex: Psychological Press.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry, 20*(1), 11-21.

Shulman, H. G. (1970). Encoding and retention of sematic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9*(5), 499-508.

Shulman, H. G. (1972). Semantic confusion errors in short-term memory. *Journal of Verbal Learning & Verbal Behavior, 11*, 221-227.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science, 14*(5), 455-461.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122-129.

Warrington, E., & Shallice, T. (1969). Selective impairment of auditory verbal short-term memory. *Brain, 92*, 885-896.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72*(2), 89-104.

Wickens, D. D. (1973). Some characteristics of word encoding. *Memory and Cognition, 1*, 485-490.

Wolters, G., & Goudsmit, J. J. (2005). Flashbulb and event memory of September 11, 2001: Consistency, confidence and age effects. *Psychological Reports, 96*(3), 605-619.

Yonelinas, A. P., & Levy, B. J. (2002). Dissociating familiarity from recollection in human recognition memory: Different rates of forgetting over short retention intervals. *Psychon Bull Rev, 9*(3), 575-582.

Chapter II

FALSE WORKING MEMORIES:
SEMANTIC DISTORTION IN A MERE 4 SECONDS.

Abstract

False memories are well-established, episodic memory phenomena: Semantically related associates are confidently and erroneously remembered as studied items. We report four experiments yielding similar effects in a working memory paradigm. Four semantically related words were retained over a brief interval. Whether or not the interval was filled with a math verification task, semantically related lures were mistakenly recognized as members of the memory set and took longer to reject than did unrelated negative probes. In a short-term recall task, semantic intrusions exceeded other errors (e.g., phonemic). Our results demonstrate false memory effects for a subspan list when a mere 4 sec was given between study and test. Such rapid semantic errors presumably result from associative processing, may be related to familiarity-based proactive interference in working memory, and are consistent with recent models that integrate short- and long-term memory processes.

Introduction

Episodic remembering is often associated with the feeling that we are, in some way, reexperiencing or reliving past events. We may vividly recall sights, sounds, tastes, and feelings associated with events that occurred minutes, days, or years ago. Despite this subjective experience, research shows that our long-term memories are not always reliable representations of the past (Roediger, 1996; Roediger & McDermott, 1995; Schacter & Slotnick, 2004). Such memories, rather, may differ dramatically from true events and, in some cases, may actually be "false," either in part or in their entirety, reflecting events that never truly occurred (Loftus, 1997; Roediger, 1996; Roediger & McDermott, 1995).

Research on false memories has focused on long-term episodic memory, exploring how recall and recognition of previously studied material can become distorted, generally over a period ranging from minutes to days. In the Deese–Roediger–McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995), participants study 15-item word lists, all of which are semantically related to an unstudied theme word. The participants then recall as many words as possible from the list or recognize studied words presented in a second list that includes studied words, the unstudied theme word, and other unassociated and unstudied foils. Investigations using the DRM paradigm consistently show that participants both falsely recall and falsely recognize unstudied theme words. Confidence ratings or remember/know judgments typically indicate moderate confidence in these decisions (e.g., Frost, 2000; Roediger & McDermott, 1995; Schacter & Slotnick, 2004).

Although the false memory phenomenon has been investigated largely in the long-term memory domain, search on short-term working memory indicates that this system too is fallible (Baddeley, 2003; Baddeley & Hitch, 1974). Working memory is susceptible to several forms of interference, wherein accurate and timely recall or recognition can be disrupted by interfering information. The interfering effects of phonological and visual similarity were established early on (Arons & Mason, 1999; Conrad, 1964; Logie, Della Sala, Wynn, & Baddeley, 2000). However, despite several well-documented semantic effects, such as semantically based proactive interference (Wickens, 1973) and semantic influences on memory span (Baddeley, 1966; Crowder, 1978; see also Hulme, Maughan, & Brown, 1991), semantic interference effects on short-term remembering were originally found to be less robust, less consistent (e.g., Baddeley, 1966, 1972; Baddeley & Dale, 1966; Baddeley & Levy, 1971; Dale & Gregory, 1966; Kintsch & Buschke, 1969; Shulman, 1970, 1972), and generally weaker than those associated with long-term memory. Consequently, semantic coding was historically considered a signature of episodic long-term memory, whereas models of verbal short-term working memory emphasized lower level phonological codes (Baddeley, 1966, 1972, 1986). This dichotomy, which is inconsistent with many current models of memory (Baddeley, 2000; Cowan, 1999, 2005; Haarmann & Usher, 2001; Nairne, 2002), may have diverted interest away from the question of whether short-term remembering is susceptible to false memory effects.

More recently, the possibility of short-term semantic coding has become more widely recognized. Patient studies indicate that working memory may include separable components dedicated to the short-term maintenance of semantic information (Hamilton

16

& Martin, 2005; Martin, Shelton, & Yafee, 1994). Moreover, the idea of short-term

semantic representations is readily handled by models that conceptualize working

memory as the activated portion of long-term memory (Cowan, 1999, 2005; Haarmann &

Usher, 2001; Nairne, 2002; Oberauer, 2002; cf. Baddeley, 2000). The view that working

memory and episodic long-term memory are mediated by overlapping and shared

components is also supported by evidence gleaned from brain imaging studies (e.g.,

Cabeza, Dolcos, Graham, & Nyberg, 2002; Karlsgodt, Shirinyan, van Ep, Cohen, &

Cannon, 2005; Ranganath, Johnson, & D'Esposito, 2003).

These developments are consistent with the possibility that false memories may

occur rapidly, on the time scale of working memory tasks, and may not require the use of

supraspan lists of memoranda. Rapidly occurring associative processes constitute the

basis for one of the popular explanations of the false memory effect in long-term

remembering (Roediger, McDermott, & Robinson, 1998). According to this account,

theme words semantically related to studied items become activated via associative

processes in semantic memory (Collins & Loftus, 1975; Underwood, 1965; see also

Arndt & Reder, 2003), much like semantic priming effects that are evident within

seconds following exposure to a prime (Meyer & Schvaneveldt, 1971).

Our present purpose, therefore, was to test the hypothesis that false recognition

and false recall can occur within a canonical working memory paradigm, with subspan

lists of memoranda. The first pair of experiments (1A and 1B) used recognition and

recall, respectively. Memory sets of four semantically related words were followed by a

3- to 4-sec retention interval that was filled with a mathematical distractor task, making

trials similar in structure to the *operation span* task that is widely used to investigate

17

working memory (Smith et al., 2001; Turner & Engle, 1989). In the recognition task (Experiment 1A), the retention interval was followed by one of three probe types: a *positive probe* that had appeared in the memory set for that trial, a *negative probe* that had not appeared and was not associated with items in the memory set, or a *lure probe* that was semantically associated but not part of the memory set. In the recall task (Experiment 1B), participants engaged in free recall of memory items following the retention interval. Identical procedures were used in the second pair of experiments (2A and 2B), except that the distractor task was omitted, thereby permitting subvocal rehearsal. In all four experiments, false memory effects were robust, indicating that semantic memory distortions can occur when attempting to retain only four items over a mere 3- to 4-sec delay.

We also predicted that response time (RT) measures of recognition performance would be sensitive to the semantic association between the probe and the memory set, consistent with previous findings by Bartha, Martin, and Jensen (1998). In an item recognition task in which four unassociated words were used as memoranda, Bartha et al. found significant slowing when participants correctly rejected nonstudied synonyms or associates of the memory set. In the present experiments, we found that participants took longer to reject lures than negative probes, revealing robust semantic interference effects.

EXPERIMENT 1A

Method

Participants

The participants were 24 right-handed University of Michigan students (12 of them male; mean age: 19.25 years). The participants in this experiment and in all the subsequent experiments participated for course credit.

Materials

The stimuli were presented on Dell personal computers using E-Prime 1.1 software (www.pstnet.com). Button-press responses were collected using a Cedrus Model RB-730 response pad (www.cedrus.com). All the stimuli appeared in a black Arial 20-point font.

Item recognition task. Forty-eight lists were created from a sub-set of 55 previously published lists shown to elicit episodic false memories (Roediger, Watson, McDermott, & Gallo, 2001). Each list consisted of four semantically related words, all associates of a common theme word. Only one memory set of four words was created from each original list. These 48 new lists were divided into four groups of 12 four-word lists (Groups A–D) that were equated in mean backward associative strength ($M = .34$ for each group of lists). Theme words served as the probes on all the trials, and no theme words were repeated for a given participant. Each participant completed 36 trials: 12 trials were paired with a lure, which was the theme word associated with that list; 12 trials were paired with an unrelated negative probe, which was a theme word associated with a nonpresented list; and 12 trials were paired with a positive probe. On positive probe trials, one item from the constructed list was replaced with the theme word for that

19

list, which also served as the positive probe. Probe type was counterbalanced with word lists across participants, so that for one quarter of all the participants, lists in Group A were paired with lure probes, lists in Group B with negative probes, and lists in Group C with positive probes. Themes associated with Group D lists served as the unrelated negative probes. There were a total of four counterbalanced groups, with $N/4$ subjects in each. This procedure ensured that all the participants encountered the same probes, all theme words, but in different contexts, as lures, negative probes, or positive probes. Trials were presented in random order for each participant. No participant was exposed to a given theme or probe more than once during the experiment.

   Distractor task. A math equation verification task based on the operation span task (Turner & Engle, 1989) was adapted from Smith et al. (2001). During the retention interval of the memory task, the participants viewed a completed math equation, decided whether or not it was solved correctly, and responded accordingly. In each equation, the first operation was always multiplication or division; the second was always addition or subtraction.

Procedure

   The participants first provided written informed consent; all research was approved by the Behavioral Sciences Institutional Review Board at the University of Michigan. A short set of practice trials, with memory sets consisting of four unrelated words, then familiarized them with the task demands. At the beginning of each trial, the participants saw a black fixation cross, which turned white 500 msec prior to the presentation of the memory set. The four memory set items, printed in lowercase font, then appeared for 1,200 msec. The retention interval varied randomly between 3,000 and

4,000 msec, during which the participants completed the math verification distractor task. The math equation was displayed for 3,000 msec, appearing randomly 0, 150, 250, or 500 msec following the offset of the memory set. The participants made a left-handed button-press response to indicate whether the math equation was solved correctly or incorrectly. A fixation cross appeared during the portion of the retention interval not occupied by the math equation (0, 150, 250, or 500 msec). Next, the probe word appeared in capital letters at the center of the screen. Using their right hand, the participants pressed one button to indicate that the probe word had appeared in the memory set, disregarding case, and another button if it had not. A black fixation cross appeared during the inter-trial interval, which varied randomly between 1,500 and 2,000 msec. The participants were instructed to respond as accurately as possible on the math verification task and to respond as quickly and accurately as possible on the item recognition task.

Results and Discussion

Mean accuracy on the math task was 0.81 ($SE = .02$). Mean probe accuracy and RTs for each probe type were compared using paired t tests. Significant values are reported after Bonferroni correction for multiple comparisons, unless otherwise noted. Effect sizes for correlated designs were computed using original standard deviations for each condition mean (Dunlop, Cortina, Vaslow, & Burke, 1996).

Our primary comparisons of interest were the accuracy and RT differences between negative probes and lures. Table 1 shows the proportion of *yes* responses for each probe type. The participants were nearly twice as likely to falsely recognize lure probes than negative probes as members of the memory set, and this difference was

significant [$t(23) = 3.43, p < .01, d = .82$]. The participants were more accurate in responding to positive (correct *yes* responses) versus lure probes (correct *no* responses) [$t(23) = 2.34, p < .05, d = .58$]. There was no accuracy difference between responses to positive and negative probes [$t(23) = 0.98$, n.s.]. Mean correct RTs for negative, lure, and positive probe trials are displayed in Figure 1A. RTs for lure probes were significantly longer than those for both negative probes [$t(23) = 5.04, p < .001, d = .70$] and positive probes [$t(23) = 4.17, p < .001, d = .75$]. No reliable differences were observed between RTs to positive and negative nonrelated probes [$t(23) = 0.30$, n.s.], indicating that the participants were equally fast in correctly rejecting negative probes and correctly recognizing positive probes.

	Probe Type	Proportion of *Yes* Responses	
		M	SE
Experiment 1A: Distraction	Positive	.80	.03
	Negative	.16	.03
	Lure	.31	.04
Experiment 2A: No Distraction	Positive	.96	.01
	Negative	.01	.01
	Lure	.09	.03

Table 2.1
Mean proportion of positive, negative and lure probes to which participants responded 'yes', indicating that the probe was recognized as a member of the memory set. Positive probes are those that did appear in the memory set. Negative probes were not members of the memory set, nor were they semantically related to memorized items. Lure probes were not present in the memory set but were semantically related to items in that set. In both experiments, lures were falsely recognized significantly more often than negative probes.

In this experiment, the three probe types occurred equally often. Consequently, completely accurate performance required negative responses on two thirds of the trials, which could lower the accuracy for positive probes. To reduce this *no* bias and to ensure that any such bias did not contribute to our false recognition effects, 18 new participants

performed a similar experiment in which there were equal numbers of *yes* and *no* trials.

Positive probe recognition was higher, and accuracy and RT differences between lures

and negative probes remained highly significant[1]. The subsequent recognition experiment

(2A) used an equal distribution of probe types because this design is more efficient for

examining false recognition errors.
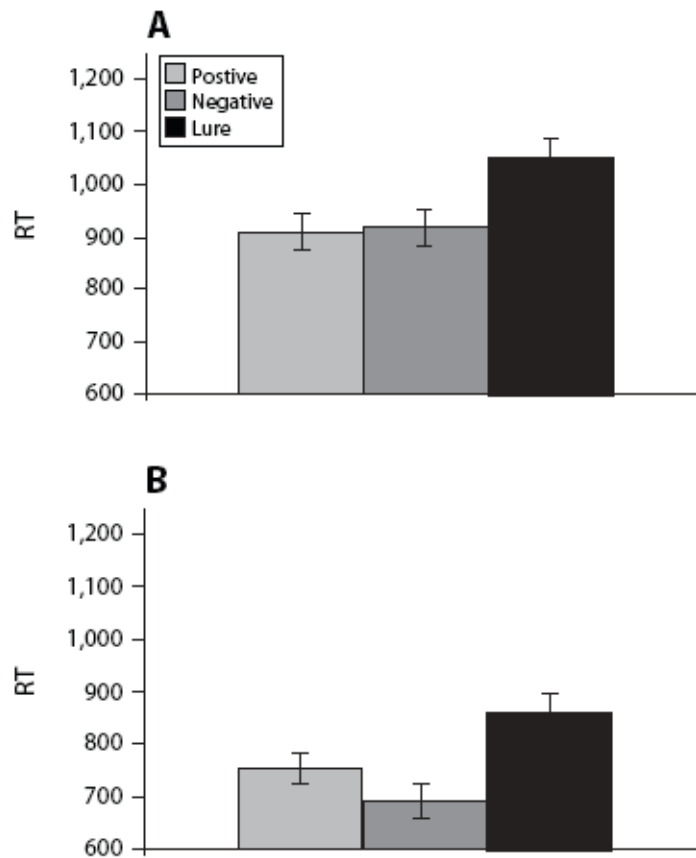
Figure 2.1
Mean response times (RTs, in milliseconds) for correct responses to positive,
negative, and lure probes in Experiment 1A, distraction (A) and Experiment 2A,
no distraction (B). The semantic interference effects, defined as the RT
difference for correct rejections of lures versus unrelated negative probes, were
134 and 166 msec for Experiments 1A and 2A, respectively.

---

[1] Mean recognition accuracy was .86 ($SE = .02$), .93 ($SE = .01$), and .77 ($SE = .04$) for positive, negative, and lure probes, respectively. Mean RT for correct trials was 796.31 msec ($SE = 30.13$) for positive probes, 810.57 msec ($SE = 30.02$) for negative probes, and 982.44 msec ($SE = 40.40$) for lure probes. Mean accuracy and RT differences between lures and negative probes were significant [$t(17) = 4.24, p < .01, d = 1.22$, and $t(17) = 7.39, p < .001, d = 1.14$, respectively].

EXPERIMENT 1B

False recognition of semantically related lures is well documented in episodic memory. Experiment 1A demonstrated that this effect can occur on the timescale of a working memory task. However, in the long-term memory domain, false memories also manifest as semantic intrusions in free recall. The following experiment demonstrated false recall in short-term remembering when there is distraction present in the retention interval.

Method

Participants

The participants were 18 new right-handed University of Michigan students (9 of them male; mean age = 19.83 years).

Materials and Procedure

This experiment used the same 48 four-item lists as those utilized in Experiment 1A, except that theme words were never presented as memoranda. No theme or memorandum was ever repeated. Each memory set appeared for 1,200 msec, followed by a 3- to 4-sec retention interval during which the participants completed the math distractor task described above. Next, a prompt appeared instructing the participants to say the words from the memory set aloud, in any order. Vocal responses were recorded directly onto the computer hard drive with the use of a standard Labtec microphone and HarddiskOgg audio capture software (Copyright 2004 Jan Lellmann/Fridgesoft). The recall period was untimed.

Results and Discussion

Recall responses were transcribed and coded manually as either correct or incorrect. Incorrect responses were further classified as follows: (1) *semantic*—that is, the word was the theme word for the trial, an associate of that theme (not presented, but listed on the original 15-item DRM list from which our 4-item list was created; Roediger et al., 2001), or a word judged by two trained coders as being related in meaning to two or more items in the memory set; (2) *phonological* (the word sounded like 1 or more items in the memory set and was not a semantic associate of the theme word); or (3) *other*, a category that included both repeated (correct or incorrect) responses and recalled words that were related in neither meaning nor sound to those in the memory set. Note that semantic errors were classified conservatively, in that words associated with only one member of a memory set and words that did not meet the criteria for any category (approximately 2% of the errors) were omitted from analysis. Nonword and unintelligible utterances (approximately 4% of the errors) were also omitted.

The participants correctly recalled an average of 3.22 words per trial ($SE = 0.10$). Mean recall errors by error type are displayed in Figure 2A. The main effect for error type was significant [$F(2,34) =15.33, p<.001, \eta2 =.47$]. The participants made more semantic than phonological recall errors [$t(17) = 4.56, p < .01, d = 1.36$]. In fact, the participants made more than twice as many semantic errors as all other error types combined, and this difference was reliable [$t(17) = 3.37, p < .01, d = 0.77$]. Of all the responses coded as semantic errors, 66% consisted of the theme word associated with the memory set, and 18% were nonpresented (semantically associated) items from the original DRM lists. Mean recall positions for semantic, phonological, and other errors

were 3.01 (*SE* =0.14), 2.46 (*SE* =0.24), and 2.75 (*SE*=0.27), respectively, which were not

significantly different. Mean math accuracy was .81 (*SE=*02).
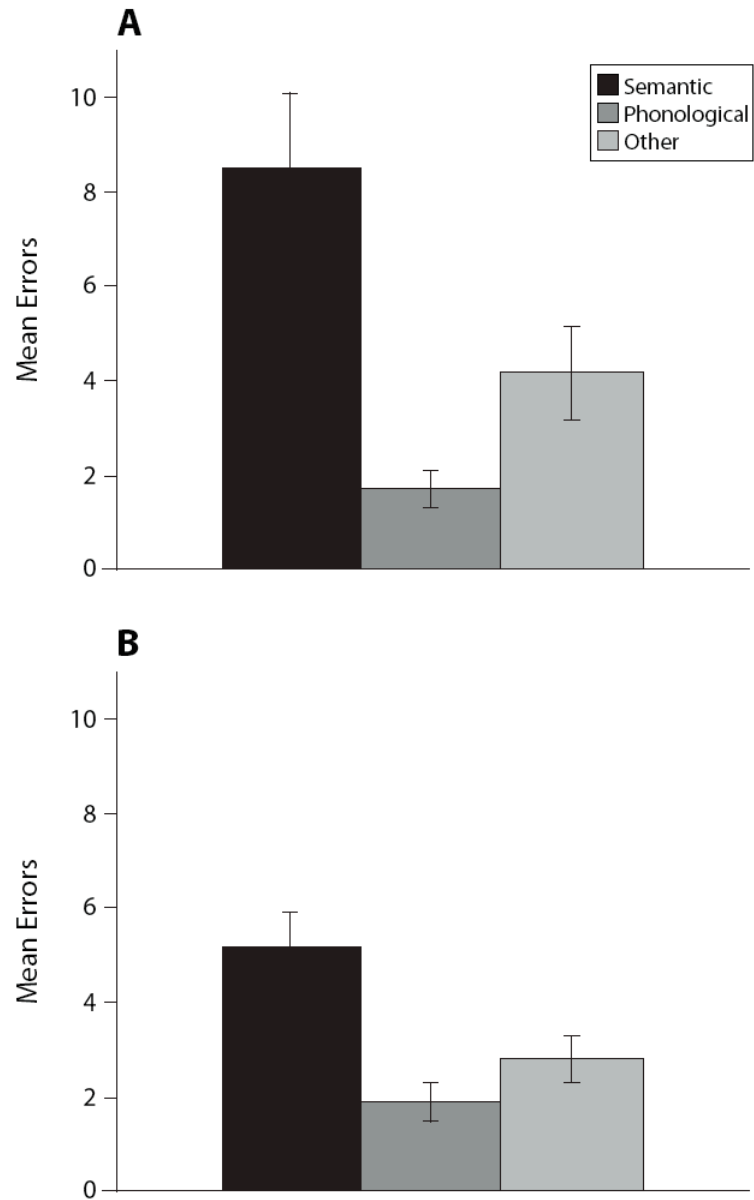
Figure 2.2
Mean number of *semantic*, *phonological*, and *other* recall errors in Experiment 1B,
distraction (A) and Experiment 2B, no distraction (B). See the text for details.

EXPERIMENT 2A

Experiments 1A and 1B demonstrate that false recognition and false recall errors can occur with a 4-item memory set retained over a 3- to 4- second interval. Furthermore, lure probes that were correctly rejected produced longer RTs than correct rejections of negative unrelated probes, consistent with the possibility of semantic interference.

These rapid semantic distortions were evident in a working memory task when the retention interval was filled with distraction. The completion of a math verification task while holding words in memory is also required in the operation span task (Turner & Engle, 1989), which is widely used as a measure of working memory capacity (see e.g., Reuter-Lorenz & Jonides, 2007). The use of this sort of distracter task during short-term retention, therefore, has a strong precedence in the working memory literature (Miyake & Shah, 1999).

However, the math verification requirement presumably interfered with subvocal rehearsal of list items, and may have encouraged the use of episodic memory processes. If true, then omitting the distracter task should permit verbatim retention of 4 items for 4 seconds and possibly eliminate the disproportional incidence of false memory errors. Experiments 2A and 2B respectively examine the extent to which false recognition and recall occur when the math verification task is omitted.   If false memories persist in the absence of distraction, then these errors cannot be simply attributed to the disruption of rote rehearsal.

Method

Participants

Twenty-five new right-handed University of Michigan students (12 of them male; mean age = 20.3 years) participated.

Materials and Procedure

The experiment was the same as Experiment 1A, except that the math verification task was omitted. To better equate the perceptual and response demands of the two experiments, a row of empty boxes appeared in place of the math problem during the retention interval; the participants were required to make a left-handed button press response to the onset of the boxes.

Results and Discussion

Mean proportions of *yes* responses for each probe type are shown in Table 1. As in Experiment 1A, there was a significant false memory effect, in that the participants were more likely to falsely recognize lure probes than negative probes as members of the memory set [$t(24) = 3.13$, $p < .01$, $d = 0.79$]. Accuracy for positive and negative probes did not differ [$t(24) < 2.24$, n.s.].

In order to further examine the effects of distraction, 2 X 2 mixed participants ANOVAs were performed to compare accuracy and RT (correct trials) for negative and lure probes with and without distraction. The results showed main effects of distraction on both probe accuracy [$F(1,47) = 29.91$, $p < .001$, $\eta2 = .39$] and RT [$F(1,47) = 19.73$, $p < .001$, $\eta2 = .30$], indicating that the presence of distraction reliably decreased accuracy and lengthened RT. There were no interactions between probe type and distraction on either

28

dependent measure, indicating that the false recognition and semantic interference effects did not reliably differ as a function of distraction.

EXPERIMENT 2B

Method

Participants

The participants were 20 new right-handed University of Michigan students (11 of them male; mean age= 18.75 years).

Materials and Procedure

The experiment was the same as Experiment 1B, except that the math verification task was omitted and replaced with the simple key press to the onset of boxes, as in Experiment 2A.

Results and Discussion

Recall responses were transcribed and coded manually as either correct or incorrect. Incorrect responses were further classified as *semantic*, *phonological*, or *other*, as described in Experiment 1B. The error-coding criteria were the same as those used in Experiment 1B. Six percent of the errors that did not meet the criteria for any error category and another 6% that were nonwords and unintelligible utterances were omitted from the analysis.

The participants correctly recalled an average of 3.51 words per trial ($SE=0.06$). Mean recall errors by error type are displayed in Figure 2B. The main effect for error type was significant [$F(2,38)=11.27, p < .001, \eta2=.37$]. The participants made more than twice as many semantic as phonological recall errors [$t(19)=3.89, p < .01, d=1.26$]. As in

Experiment 1B, the participants made more semantic errors than all other error types combined, and this difference was reliable [$t(19) = 3.01$, $p < .05$, $d = 0.86$]. Of all the responses coded as semantic errors, 46% consisted of the theme word associated with the memory set, and 24% were nonpresented items from the original DRM lists (Roediger et al., 2001). Mean recall positions for semantic, phonological, and other errors were 2.93 ($SE = 0.16$), 2.34 ($SE = 0.23$), and 2.76 ($SE = 0.21$), respectively, and these did not differ reliably.

A *post hoc* independent samples *t* test compared recall accuracy with and without distraction. As was expected, recall accuracy was higher when no distraction was presented during the retention interval [$t=2.56$, $p < .05$, $d=0.82$]. In order to further examine the effects of distraction, a 2 X 2 mixed participants ANOVA compared the rate of recall errors (semantic, phonological, and other) with and without distraction. The results showed a main effect of error type [$F(2,72)=26.95$, $p < .01$, $\eta2=.43$], reflecting the prevalence of semantic errors, and a marginal error type x distraction interaction [$F(2,72)=3.11$, $p=.051$, $\eta2=.08$]. Examination of the means (Figure 2B) suggests a trend toward an increase in semantic and other errors in the presence of distraction and a relatively stable rate of phonological errors.

GENERAL DISCUSSION

In the present research, we investigated the occurrence of false memory effects in the working memory domain and demonstrated two robust phenomena: First, semantically related lures were falsely recognized and falsely recalled a mere 3–4 sec following study. False recognition effects occurred reliably whether or not the retention

30

interval was filled with a distracting task. The recall results from Experiment 2B are particularly striking, in that semantic intrusions were prevalent after a brief retention interval during which the participants were free to engage in subvocal rehearsal. These false recall errors indicate that short-term memory distortions do not require the presentation of a semantically associated lure probe or the completion of a distractor task during the retention interval.

Second, salient interference is evident in the RTs for correct rejections of semantically related probes. As compared with unrelated probes, semantic associates took well over 100 msec longer to reject, whether the retention interval was filled or unfilled.

The present research introduces a procedure for generating reliable false memory errors within the temporal and set size parameters characteristic of working memory. The use of four associated memoranda distinguishes this work from previous investigations in which false recognition effects were less robust. For example, in a relatively recent item recognition study, Bartha et al. (1998) used memory sets of four unassociated items. As compared with the present results, the false recognition effects in Bartha et al. were weaker (Experiment 3) and did not replicate in a second experiment (Experiment 4).

To what extent might the present results be due to strategic processes that participants adopt to contend with task demands? In recognition (Experiments 1A and 2A), might participants adopt a liberal response bias to lure probes, as in Miller and Wolford's (1999; cf. Roediger & McDermott, 1999) account of false recognition effects in the long-term memory domain? This explanation assumes that participants shift their response criteria on a trial-by-trial basis, an assumption that does not hold for all circumstances (e.g., Wixted & Stretch, 2000) but may be more likely with only a single

31

recognition test per list. However, in earlier versions of our short-term recognition task (Atkins & Reuter-Lorenz, 2005), we presented three sets of four different associates from the same original DRM lists, thus probing the same list in consecutive test blocks. The false memory effect in these versions of the task was as robust as the effects we report here. These observations cannot definitively rule out the possibility that criterion shifts contribute to the rapid false recognition effects reported here; however, they do weaken this account as the sole source of the memory illusion we observe.

A strategy-based explanation of our false recall results also deserves consideration. In particular, participants could be treating the free recall task (Experiments 1B and 2B) like a forced recall task in which they are *required* to generate four responses on every trial. False recall in the long-term memory domain is greater for forced than for free recall (see, e.g., McKelvie, 2001; Meade & Roediger, 2006). Our instructions, however, did not emphasize the requirement to recall four items, nor did the participants feel compelled to do so, as is evident from their average recall of 3.50 and 3.75 items in Experiments 1A and 2A, respectively. Nevertheless, when the participants remembered fewer than four items, their metaknowledge about the relatedness of the lists may have led them to interrogate semantic memory in search of an appropriate response. This strategy would predict that semantic errors would be most likely in the fourth position. Like other recall errors, the average output position for semantic errors was the third (2.92 and 3.01 in Experiments 1A and 2A, respectively). These considerations lead us to believe that the false memory effects we observed reveal memory's extreme vulnerability to distortion.

What mechanisms, then, might underlie the robust false memory errors that we report? Three distinct hypotheses can be entertained. Common to all of them is the assumption that long-term semantic memory is the basic source of associative processing that mediates activation of thematically related lures and associates (Roediger et al., 2001). The accounts differ with respect to whether the mechanisms that access semantic codes are (1) intrinsic to working memory, (2) intrinsic to episodic memory, or (3) shared and common to both short- and long-term remembering.

Despite controversy as to whether there exists a working memory system with independent semantic components (e.g., Baddeley, 2003; Haarmann & Usher, 2001; Hamilton & Martin, 2005; Shivde & Thompson-Schill, 2004), support for such a system comes from behavioral evidence (e.g., Potter, 1993), as well as from neuropsychological and neuroimaging results (e.g., Martin et al., 1994; Shivde & Thompson-Schill, 2004), indicating that word meanings can be retained over the short-term by working memory processes that are independent from phonological and visual codes. According to this model, the semantic errors in the present paradigm would arise from similarity-based interference during the encoding, maintenance, or retrieval of semantic codes, using a working memory subsystem that is specialized for semantics. A somewhat weaker form of this hypothesis may not require an entirely separable semantic working memory subsystem. For example, the episodic buffer, as proposed by Baddeley (2000), is a slave system to working memory's central executive that is separate from long-term memory and that contains multidimensional representations of information gathered from several sources, including long-term episodic memory, as well as the phonological and visuospatial buffers.

The second hypothesis follows from the view that although semantic errors in the present paradigm occur in the context of a working memory task, deeper semantic coding is the sole province of episodic memory processes that are operating in parallel with working memory. According to this view, short- and long-term false memories differ only in time course and are mediated by identical psychological processes. Phonological representations of the memoranda are maintained in working memory; however, false memories arise from episodic memory processes that are responsible for generating semantic distortions at longer delays (i.e., associative and monitoring processes operating during encoding and retrieval; see, e.g., Roediger et al., 2001). According to this view, the false memories that we observed are a product of long-term memory contributions to working memory.

The third hypothesis is that the false memory errors that we observe arise from processes that are shared by both short- and long-term remembering (see, e.g., Nairne, 2002, for a review). The model proposed by Cowan (1999) fits best in this class of hypotheses, in that long-term memory is postulated to provide the representational basis for short-term remembering. According to this view, short-term memory provides the activation processes that select and maintain the relevant portion of long-term memory, in a manner resembling extended priming. The unitary model proposed by Nairne (e.g., 2002), in which short-term memory, like episodic memory, is cue based, also falls into this third type of account. According to Nairne's feature model, what is retained over the short or long-term is not a set of particular items but, rather, a constellation of cues that include phonological, lexical, and semantic features. At retrieval, these cues are used to

34

reconstruct the memoranda. Accurate memory requires distinctive cues, thereby leading to the detrimental effects of phonological, as well as semantic, similarity.

Determining which of these hypotheses offers the best account of the false memory errors we observe will depend on future research. We note, however, that the present results demonstrate a striking parallel in the memory distortions that can affect both recognition and recall over the short or long-term. This convergence is consistent with the growing body of evidence indicating overlap between short- and long-term memory mechanisms (Cabeza et al., 2002; Karlsgodt et al., 2005; Ranganath et al., 2003; see also Baddeley, 2000; Cowan, 1999). Theoretical models that may best account for our findings, therefore, are ones that emphasize an integration of short- and long-term memory processes (the third type of hypothesis outlined above).

A second theoretical question posed by the present research relates to the psychological mechanisms underlying the observed interference effect. Lures took longer to correctly reject than did negative probes. The semantic relatedness of the lure item to the memory set presumably necessitated recruitment of additional processes, thereby increasing RT. What might these additional processes be? One possibility is that semantic relatedness increased the familiarity of the lure item, making it more difficult to reject than an item that had not been recently active in working memory. This need to adjudicate between an item's heightened familiarity and its status as a member of the current memory set is very similar to explanations offered to account for familiarity-based proactive interference effects that occur across trials in short-term item recognition paradigms (Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Monsell, 1978). The increased RTs for lures could also stem from invoking source memory to decide

whether the lure was actually presented or whether it was merely thought of due to strong associations with the memory set (e.g., Johnson, Hashtroudi, & Lindsay, 1993). Source memory and the need for increased contextual coding have also been linked to familiarity-based proactive interference in short-term item recognition tasks (Badre & Wagner, 2005; Jonides & Nee, 2006).

We are currently using fMRI to investigate the possibility that common prefrontal mechanisms mediate semantic and familiarity-based interference effects (Atkins, Walsh, & Reuter-Lorenz, 2006). This ongoing neuroimaging research may also prove indispensable in distinguishing among the theoretical accounts of false working memory, in that it will allow us to examine the overlap and dissociations between the neural correlates of false short- and long-term remembering.

Taken together, our findings present strong evidence that semantic distortions of memory can occur over the very short-term and reveal a surprising fallibility of verbatim memory processes. In mere seconds following presentation of a small memory set, participants falsely recall and recognize semantically related lures. Because the contents of working memory figure prominently in reasoning and problem-solving abilities, rapid distortions of meaning that go unchecked could have considerable impact on the quality of our decisions and the accuracy of our solutions. From a theoretical perspective, the data we report converge with a number of recent findings demonstrating considerable overlap in the operations and mechanisms of episodic and working memory (e.g., Cabeza et al., 2002; Nairne, 2002; Ranganath et al., 2003; Wagner, 1999; see also, Karlsgodt et al., 2005). False recognition and false recall may, therefore, constitute a valuable new

province in which to investigate the relationship between short- and long-term

remembering.

ACKNOWLEDGEMENTS

REFERENCES

Arndt, J. & Reder, L.M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory and Language, 48,* 1-15.

Arons, S. E. & Mason, A. (1999). Effects of visual similarity on serial report and item recognition. *seconds. Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 52*, 217-240.

Atkins, A. S. & Reuter-Lorenz, P. A. (2005). 'False' working memories: Memory distortion in a mere 4 seconds. *Abstracts of the Psychonomic Society, 10*, 28.

Atkins, A. S., Walsh, M.K. & Reuter-Lorenz, P. A. (2006). False working memory: Rapid memory distortion and its neural correlates. *Journal of Cognitive Neuroscience, Supplement, 18*, 140.

Baddeley, A. D. (1966). Short term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18A*, 362-365.

Baddeley, A.D. (1972). Retrieval rules and semantic coding in short-term memory. *Psychological Bulletin, 78*, 379-385.

Baddeley, A. D. (1986). *Working Memory*. Clarendon: Oxford University Press.

Baddeley, A.D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science*, 1, 417-423.

Baddeley, A. D. (2003). Working memory: Looking forward, looking back. *Nature Neuroscience Reviews, 4*, 829-839.

Baddeley, A. D. & Dale, H. C. A. (1966). The effects of semantic similarity on retroactive interference in long and short term memory. *Journal of Verbal Learning and Verbal Behavior, 5*, 417-420.

Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *Recent Advances in Learning and Motivation (Volume 8)*. New York: Academic Press.

Baddeley, A. D. & Levy, B. A. (1971). Semantic coding and short-term memory. *Journal of Experimental Psychology, 89*, 132–136.

Badre, D. & Wagner, A. D. (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex, 15*, 2003-2012.

Bartha, M. C., Martin, R. C. & Jensen, C. R. (1998). Multiple interference effects in short-term recognition memory. *American Journal of Psychology, 111*, 89-118.

Cabeza, R., Dolcos, F., Graham, R. & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage*, 16, 317-330.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55*, 75-84.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (eds.), *Models of Working Memory: Mechanisms of active maintenance and executive control.* Cambridge, U.K.: Cambridge University Press. (pp. 62-101).

Cowan, N. (2005). *Working memory capacity*. New York, NY: Psychology Press.

Crowder (1978). Memory for phonologically uniform lists. *Journal of Verbal Leaning and Verbal Behavior, 17*, 73-89.

Dale, H. C. A. & Gregory, M. (1966). Evidence of semantic coding in short-term memory. *Psychonomic Science*, 75-76.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58, 17-22.

Dunlop, W. P., Cortina, J. M., Vaslow, J. B. & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1*, 170-177.

Frost, P. (2000). The quality of false memory over time: is memory for misinformation "remembered" or "known"? *Psychonomic Bulletin & Review, 7*, 531-536.

Haarman, H. & Usher, M. (2001). Maintenance of semantic information in capacity limited short-term memory. *Psychonomic Bulletin and Review, 8*, 568-578.

Hamilton, A. C. & Martin, R. C. (2005). Dissociations among tasks involving inhibition: A single-case study. *Cognitive Affective & Behavioral Neuroscience, 5*, 1-13.

Hulme C., Maughan, S. & Brown, G.D.A. (1991). Memory for familiar and unfamiliar words: evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*, 685-701.

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

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139*, 181-193.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A. & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 8410-8413.

Karlsgodt, K. H., Shirinyan, D., van Ep, T. G. M., Cohen, M. S. & Cannon, T. D. (2005). Hippocampal activations during encoding and retrieval in a verbal working memory paradigm. *Neuroimage, 25,* 1224-1231.

Kintsch, W. & Buschke, H. (1969). Homophones and synonyms in short-term memory. *Journal of Experimental Psychology, 80*(3P1), 403-407.

Loftus, E. F. (1997). Creating false memories. *Scientific American, 277*, 70-75.

Logie, R. H., Della Sala, S., Wynn, V. & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology, 53*, 626-646.

Martin, R. C., Shelton, J. R. & Yafee, L. S. (1994). Language processing and working memory: Evidence for separate phonological and semantic capacities. *Journal of Memory & Language, 33*, 83-111.

McKelvie, S. J. (2001). Effects of free and forced retrieval instructions on false recall and recognition. *Journal of General Psychology, 128*, 261-278.

Meade, M. L., & Roediger, H. L. (2006). The effect of forced recall on illusory recollection in younger and older adults. *American Journal of Psychology, 119,* 433-462.

Meyer, D. E. & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.

Miller, M.B. & Wolford, G.L. (1999). The role of criterion shift in false memory. *Psychological Review, 106*, 398–405.

Miyake, A. & Shah, P. (Eds.) (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Monsell, S. (1978). Recency, immediate recognition memory, and reaction-time. *Cognitive Psychology, 10*, 465-501.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53-81.

Oberauer, K. (2002). Access to information in working memory. Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 28, 411-421.

Potter, M. C. (1993). Very short-term conceptual memory. *Memory & Cognition, 21*, 156-161.

Ranganath, C., Johnson, M.K. & D'Esposito, M. (2003). Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia*, 41, 378-389.

Reuter-Lorenz, P. A. & Jonides, J. (2007). The executive is central to working memory: Insights from age, performance and task variations. In Conway,A.R., Jarrold, C., Kane, M.J., Miyake, A., Towse, J.N., (Eds.), *Variations in Working Memory*: Oxford University Press, pages 250-270.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language, 35*, 76-100.

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

Roediger, H.L. & McDermott, K.B. (1999). False alarms about false memory. *Psychological Review, 106*, 406–410.

Roediger, H.L., McDermott, K.B. & Robinson, K.J. (1998). The role of associative processes in producing false memories. In M. Conway, S. Gathercole & C. Cornoldi (Eds.), *Theories of Memory II* (pp. 187-245). Hove, Sussex: Psychological Press. .

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

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

Shivde, G. & Thompson-Schill, S. (2004). Dissociating semantic and phonological maintenance using fMRI. *Cognitive Affective and Behavioral Neuroscience, 4*, 10-19.

Shulman, H. G. (1970). Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9*, 499-508.

Shulman, H. G. (1972). Semantic confusion errors in short-term memory. *Journal of Verbal Learning & Verbal Behavior, 11*, 221-227.

Smith, E. E., Geva, A., Jonides, J., Miller, A., Reuter-Lorenz, P. & Koeppe, R. A. (2001). The neural basis of task-switching in working memory: Effects of performance and aging. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 2095-2100.

Turner, M. L. & Engle, R. W. (1989). Is working memory capacity task dependent. *Journal of Memory and Language, 28*, 127-154.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology, 70*, 122-129.

Wagner, A.D. (1999). Working memory contributions to human learning and remembering. *Neuron, 22*, 19-22.

Wickens, D.D. (1973). Some characteristics of word encoding. *Memory and Cognition, 1*, 485-490.

Wixted, J. T., & Stretch, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review, 107*, 368-376.

Chapter III

NEURAL MECHANISMS OF SEMANTIC INTERFERENCE AND VULNERABILITY TO FALSE RECOGNITION IN SHORT-TERM MEMORY

Introduction

Distortions of memory have been a subject of interest for cognitive psychology since its inception. One reason for this is that examination of the circumstances under which our memories fail us can illuminate our understanding of how memory is organized. In the last two decades, the term *false memory* has come to describe instances in which episodic memories become distorted, leading to false recognition and recall of previously unstudied items. In the Deese-Roediger-McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) paradigm, participants study lists of 12 -15 words which are all related in meaning to a common unstudied theme word, or lure. At test, participants are required to either recognize studied items from a list of probes that includes unstudied lure words, or to recall studied items in free report. Investigations using variants of this procedure have shown that participants consistently and confidently recognize unstudied lures, and even produce these items in free recall (see Gallo, 2006 for review).

Although initial investigations of the false memory phenomena were limited to paradigms which included long study lists and retention intervals that varied from several seconds to many hours, there is recent evidence that false memories are produced rapidly, within the time and load constraints of traditionally defined short-term memory tasks (Atkins & Reuter-Lorenz, 2005; Coane, McBride, Raulerson, & Jordan, 2007; Atkins & Reuter-Lorenz, 2008). For example, using a short-term variation of the DRM (ST-DRM) paradigm, we recently demonstrated reliable false recognition and recall of unstudied lures only 4 seconds following encoding (Atkins & Reuter-Lorenz, 2008). Furthermore, in the recognition version of our task, we found strong evidence that the semantic relationship between lure probes and memoranda induced interference even when lures were not falsely recognized.

The increased time required to reject lure probes is consistent with the notion that correct rejection of these items requires engagement in a controlled process which can resolve interference induced by the semantic familiarity of these items. When interference resolution is successful, the lure probe can be correctly rejected. False recognition of these items could indicate either a failure of this controlled process, or a failure to engage it at all.

Interestingly, false recognition has not been widely investigated as a failure of cognitive control processes. One reason for this is the paucity of crosstalk between those investigating false memory in the episodic domain, and those investigating interference resolution in short-term memory. Interference resolution processes in short-term memory have been studied extensively using the recent probes (RP) task (Monsell, 1978; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998). In this task, participants study a set

of 4-6 memoranda. Following a brief retention interval, a probe item is presented that requires a Y/N recognition response. Generally, 4 probes types are employed. Items that require a 'No" response include *recent negative* (RN) probes that are not present on the current trial, but were members in the memory set on the trial immediately preceding the current one, and *non-recent negative* (NRN) probes that are not members of the current set and have not appeared for the last several trials. Items requiring a 'Yes' response include standard *positive* (POS) probes are members of the current memory set, and *recent positive* (RPOS) that appeared as memoranda on both the current and immediately preceding trial.

Behaviorally, findings from RP tasks consistently show that participants are markedly slower in rejecting RN relative to NRN probes, suggesting that the episodic familiarity of the RN probes makes it more difficult to correctly reject. The source of this slowing has been interpreted as the need to engage in an interference resolution process by which the temporal familiarity of the RN item can be adjudicated in the service of accurate recognition memory. The neural circuitry subserving this interference resolution has been a topic of numerous neuroimaging investigations that have high-lighted an important role for left ventrolateral prefrontal cortex (L VLPFC) in this process. Probe-related increases for RN relative to NRN probes have been shown and replicated across a variety of RP tasks (see Jonides & Nee, 2006 for review). Furthermore, behavioral indices of proactive interference (PI) in this task, calculated as RT difference between correct responses to RN relative to NRN probes, have shown positive correlations with increases in L VLPFC activity (Badre & Wagner, 2005; Jonides & Nee, 2006; Badre &

46

Wagner, 2007), indicating that increased interference is associated with increased activity in this region.

The importance of L VLPFC in familiarity-based interference is further demonstrated by findings examining age differences in PI in short-term memory, and by patient studies. For instance, Jonides et al. (2000) showed that older subjects who showed increased PI effects failed to exhibit increased activation of left BA45. Patients with focal lesions to this region also show increased PI (Thompson-Schill, et al., 1999, see also Hamilton & Martin, 2005), suggesting that engagement of L VLPFC may be necessary for the timely resolution of this form of interference.

Increased activity in L VLPFC regions has also been linked to interference resolution in paradigms other than the RP task, including tasks requiring semantic selection (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Hirshorn & Thompson-Schill, 2004; see also Poldrack, Wagner, Prull, Desmond, et al., 1999; Gold & Buckner, 2002). In the Verb Generate task, for instance, participants are asked to mentally generate a verb corresponding to a noun which is presented to them in the scanner. For example, given the noun SCISSORS, a participant may generate the response 'CUT'. Results from such investigations have shown increased L VLPFC activity and increased RT associated with the requirement to generate verbs in response to nouns that have many associated verbs (for example, 'BALL') than those that have few (for example 'SCISSORS'; Thompson-Schill, et al., 1997; Nelson, Reuter-Lorenz, Persson, Sylvester, & Jonides, 2009).

Thus, converging evidence suggests that regions of L VLPFC play a role in familiarity-induced interference and in semantic selection processes. It follows that this

47

region may also mediate semantic interference in our ST-DRM task. In this paradigm, correct rejection of lure probes requires participants to overcome the interference induced by semantic, rather than temporal familiarity, as in the RP task.

A primary goal of the present study is to investigate the role of L VLPFC in SI, and to determine the extent to which it can be dissociated from other forms of familiarity-based interference. One interpretation of L VLPFC's role in interference resolution across tasks is that this region selects between activated representations in memory (Thompson-Schill, et al., 1997). By one account, this selection process weighs evidence in supporting possible responses by comparing activations of stimulus attributes relevant to task performance (Badre & Wagner, 2007). Relevant stimulus features will vary from task to task, and could included semantic, phonological and episodic contextual cues. The greater the overlap between the features activated by probes in an item-recognition task and memorized items, the longer the selection process should take. For instance, RN probes, which have become active due to recent study, will be more difficult to reject than NRN probes that have less residual episodic activation, and therefore will induce increased L VLPFC activation. Similarly, lure probes in our ST-DRM task should show increased L VLPFC activation relative to unrelated negative probes due to the overlap of the semantic properties of these lures and memoranda.

A second aim of the present investigation is to examine veridical versus false retrieval from short-term memory. By examining cases in which interference resolution is unsuccessful (i.e. when lures are falsely recognized) we may be able to assess the extent to which regions including L VLPFC underlie these effects. One possibility is that false recognition will be associated with a failure to activate L VLPFC. Another

possibility is that L VLPFC may interact with other regions that mediate interference indexed by the L VLPFC in service of accurate memory. In this case, L VLPFC may not distinguish between trials in which interference is ultimately overcome (correct rejections of lure) from those in which it is not (false alarms to lures), but regions that support accurate recognition in the face of interference should. What might these additional regions be? One candidate proposed by Badre and Wagner (2005) is left frontopolar cortex (FPC), a region associated with episodic retrieval that showed increased activation associated with correct recognition in the face of PI in their task. Another candidate is the right posterior parietal cortex, which has shown increased activity in response to correct rejection of lure vs. unrelated probes in an episodic DRM paradigm (McDermott, Jones, Petersen, Lageman, & Roediger, 2000), and which is associated with source memory, attentional maintenance of task goals, and memory retrieval in the face of distraction (Cabeza, Ciaramelli, Olson, & Moscovitch, 2008; Ciaramelli, Grady, & Moscovitch, 2008; Singh-Curry & Husain, 2009).

The present experiment was designed to distinguish between these possibilities, and to assess the neural mechanisms associated with SI and memory distortions in the ST-DRM paradigm. We first identify and discuss regions associated with interference and veridical vs. false retrieval from short-term memory using whole-brain assessments of probe-related differences between trial types in our ST-DRM paradigm. Next, we examine brain-behavior relationships in frontal and parietal regions of interest (ROIs), and suggest a distinction between regions that respond to SI and those that support accurate task performance in the face of this interference.

49

Method

Participants

Twenty participants (12 females; mean age=20) were recruited from the University of Michigan. All participants gave informed consent as reviewed by the university's Institutional Review Board. Participants were paid $20 per hour for their participation.

Task and Procedure

Participants completed the short-term DRM paradigm (ST-DRM; Atkins & Reuter-Lorenz, 2008) during 9 task runs (Figure 3.1). At the beginning of each trial, a blinking red fixation appeared for 500ms to warn the participant the trial was beginning. This was followed by a memory set consisting of 4 semantically related items, all associated with a common theme word. The memory set appeared for 1200ms. Five hundred ms following the offset of the memory items, a dual-operation math equation appeared at the center of the screen. This equations was solved either correctly for example, $(4 \times 3) - 2 = 10$?, or incorrectly, and participants made a left-handed response to indicate whether the math was correct or incorrect. Five hundred ms following the offset of the math problem, a memory probe appeared and participants made a right-handed Y/N response indicating whether or not the probe was a member of the memory set.

During this task, theme words served as the probes on all the trials. There were

two variations of "No" trials. The first were *unrelated negative* (NEG) trials, in which

the probe consisted of the theme word associated with a nonpresented list. The second

were *lure* (LURE) trials, in which the probe consisted of the (unstudied) theme associated

with the present memory set. On *positive* (POS) trials, the associated theme was

embedded in the memory set, and served as the positive probe.
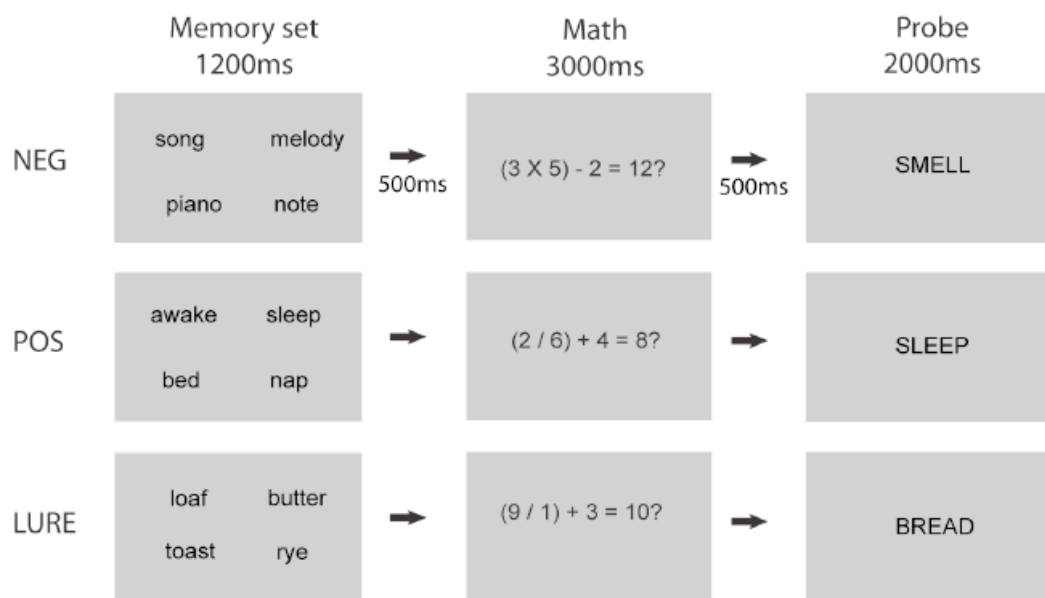
Figure 3.1.
Example trials from the ST-DRM task. Positive probes (POS) are those that did appear in the
memory set. LURE probes did not appear, but were semantically associated theme words related
to items in the current memory set. Unrelated negative (NEG) probes did not appear, and were
unrelated in meaning to items in the memory set.

With the exception of positive probes, no participant was exposed to any theme or

memoranda more than once during the experiment. Backward associative strength

(BAS), a measure of the degree of association between theme words and memoranda (see

Roediger, Watson, McDermott, & Gallo, 2001; Hancock & Hicks, 2002), was equated

across memory lists associated with each probe type, and probe type was counterbalanced

with lists across participants. This procedure ensured that participants encountered the same probes, all theme words, but in different contexts, as lures, unrelated negative probes, or positive probes. Trials were presented in random order for each participant.

Participants completed 102 ST-DRM trials which were distributed across 6 task runs. Trials were equally distributed across all three probe-types in each run. Because we were interested in examining responses to false alarms to lure probes in addition to hits and correct rejections, we used a long 16sec ITI to allow for the hemodynamic response to return to baseline between trials (Glover, 1999). All participants completed 2 practice runs prior to entering the scanner, in order to become familiar with task and response demands.

FMRI Data Acquisition

Our data were collected using a 3-T GE whole-body scanner equipped with a standard quadrature headcoil. A spiral sequence was used to collect Functional T2* blood oxygenation level-dependent (BOLD) images with 25 contiguous axial 5-mm slices. Our repetition time (TR) = 2000 ms, flip angle = 90, and field of view (FOV) = 24 cm. T1-weighted gradient echo (GRE) anatomical image was also acquired in the same FOV and slices as were used in the functional data collection. A high-resolution (60 slice) set of anatomical images was acquired via spoiled gradient-recalled acquisition in steady state (SPGR) imaging.

FMRI Data Analysis

Neuroimaging analyses were conducted using the General Linear Model implemented in SPM5 with separate regressors for each trial type in each run. Functional

and anatomical images were normalized to MNI space, and functional images were smoothed (8mm Gaussian kernel). Event-related activity to probes was modeled by convolving probe onsets with the canonical HRF. Statistical models were estimated for each participant. For each comparison of interest, contrast maps for each participant were submitted to random effects comparisons.

Behavioral Results

Behavioral findings replicate the SI and false memory effects demonstrated previously (Atkins & Reuter-Lorenz, 2008). Mean accuracy and response time (correct trials only) measures were submitted to a repeated measures analysis of variance (ANOVA). There were main effects of probe type (positive, negative, lure) on both accuracy, $F=16.98$, $p<.001$, $\eta2=.47$ and RT, $F=37.22$, $p<.001$, $\eta2=.66$. Post-hoc tests were conducted to examine false memory and semantic interference effects, and were submitted to a Bonferonni correction for multiple comparisons.

Table 3.1 shows the proportion of items recognized for each probe type. We found a reliable false memory effect, with participants falsely recognizing lure items at a rate over four times that for negative unrelated probes, $t=6.01$, $p<.001$, $d=1.54$. Mean RTs for both correct trials and false alarms (FAs) to lure items are displayed in Fig. 3.2. Participants were reliably slower to correctly reject lure probes, as compared to negative unrelated probes, $t=8.10$, $p<.001$, $d=.86$. Our SI index, measured as the difference in RT for correct rejections of lure vs. negative probes had a mean of 156.92ms (SE=19.36).

Table 3.1.

Mean proportion of positive, negative, and lure probes to which participants responded 'Yes'.

Probe Type	Proportion of *Yes* responses	
	M	*SE*
POS	.89	.02
NEG	.03	.01
LURE	.13	.02

Note- A 'Yes' response indicates that the probe was recognized as a member of the current memory set. The proportion recognized therefore represents the hit rate for POS probes, and the false recognition rate for NEG and LURE probes. The mean false memory rate, defined as the difference between false recognition for NEG and LURE probes was .10 (SE=.01).

We used paired-tests to compare RTs associated false recognition (false alarms to lures) to those associated with true recognition (hits to positive probes). False recognition of lures was reliably slower than true recognition (t=4.40, p<.001, *d*=1.02). Furthermore, RTs associated with false recognition vs. correct rejections of lures did not differ reliably, indicating that false recognition did not occur as a consequence of fast responding.
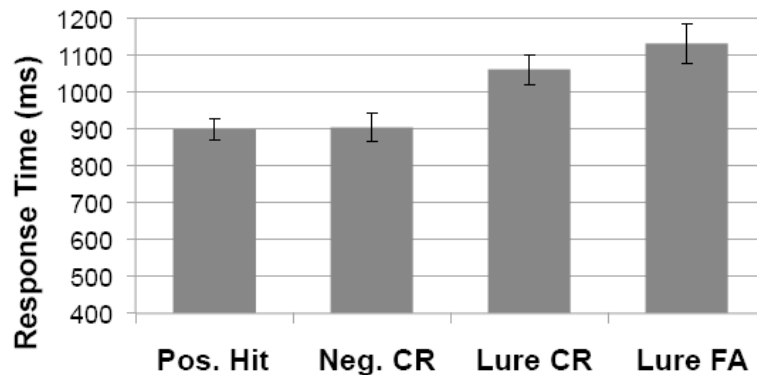

Figure 3.2.
Mean response times for correct recognition of Positive probes (hits), correct rejections (CR) of Negative and Lure probes, and false alarms (FA) to Lure probes. The mean semantic interference effect, defined as the difference in time required to correctly reject lure vs. negative probes (Lure CR - Neg.CR), was 157ms (SE=19.37).

54

Neuroimaging Results

<u>Whole-Brain Analysis</u>

Results from our whole-brain analyses are presented in Table 2, and summarized below. Unless otherwise stated all comparisons reported were significant at p<.005, uncorrected, with threshold requirement of 20 or more contiguous voxels (Forman, Cohen, Fitzgerald, Eddy, et al., 1995).

In order to examine the neural mechanisms associated with SI in the ST-DRM task, we identified regions that showed increased probe-related activity in response to correctly rejected lure probes, relative to correctly rejected negative probes. In both cases, correct 'No' responses are made to unstudied items, but in the case of a lure probe SI is present, whereas in the case of a negative unrelated probe, no interference should be present. This comparison is thus directly analogous to the recent vs. non-recent negative probe comparisons used in investigations of PI using the RP task.

Figure 3.3 displays regions that showed activity that was greater for correct rejections of lure vs. negative probes. These included a large cluster of voxels in left mid-VLPFC, with a single peak in BA45. This suggests, consistent with predictions, that L VLPFC is recruited in response to interference from unstudied items that are semantically associated with items in memory. Bilateral anterior cingulate cortex (ACC, BA24/32) also distinguished lures from negative probes. Given strong evidence associating increased ACC activity with response-level conflict across a variety of tasks

Table 3.2.
Peak voxels for whole-brain analyses (p<.005, uncorrected, 20 contiguous voxels).

	Semantic Interference (Lure CR- Neg. CR)						
	Peak			Voxels	T value	BA	Region
	x	y	z				
Frontal	-48	21	21	316	6.08	45	Left ventrolateral prefronal cortex
	-31	55	6	22	3.39	10	Left anterior prefrontal cortex
Parietal	-34	-65	48	88	5.24	40/7	Left intraparietal sulcus
	38	-58	48	32	4.61	40/7	Right intraparietal sulcus
Other	-7	21	48	182	5.22	32	Left anterior cingulate cortex
	10	31	10	56	4.01	24/32	Right anterior cingulate cortex

	True Recognition (Pos. Hit > Neg. CR)						
	Peak			Voxels	T value	BA	Region
	x	y	z				
Frontal	-31	55	6	118	5.08	10/46	Left anterior/dorsolateral prefrontal cortex
Parietal	-45	-52	48	228	5.34	40	Left inferior parietal cortex
	48	-48	45	141	5.67	40	Right inferior parietal cortex
	-10	-69	54	69	4.71	7	Left precuneus/superior parietal cortex
	17	-72	48	87	4.37	7	Right precuneus/superior parietal cortex
Other	-10	-7	21	113	4.91		Caudate

	False Recognition (Lure FA > Neg. CR)						
	Peak			Voxels	T value	BA	Region
	x	y	z				
Frontal	-10	28	36	272	6.09	9	Left dorsolateral prefronal cortex
	28	-31	60	67	3.67	4/6	Right middle frontal gyrus
	-28	0	48	48	3.8	6	Left middle frontal gyrus
	10	45	-15	32	4.15	10	Right anterior prefrontal cortex
Parietal	38	-83	36	31	5.25	7	Right precuneus/superior parietal cortex
	55	-48	42	24	4.12	40	Right inferior parietal cortex
Temporal	-65	-21	9	30	4.78	42	Left superior temporal gyrus
Other	0	-38	45	423	6.74	31	posterior cingulate
	17	45	12	53	3.81	32	anterior cingulate

	True Recogntion> False Recogntion (Pos. Hit>Lure FA)						
	Peak			Voxels	T value	BA	Region
	x	y	z				
Frontal	52	7	6	47	4.03	44	Right ventrolateral prefrontal cortex
Frontal/Partietal	-55	-14	15	45	3.84	43	Left Rolandic operculem/postcentral gyrus
Occipital	-41	-55	-21	30	5.32	37	Left fusiform gyrus
Other	-31	10	0	231	5.85		Left putamen

(e.g. Nelson, Reuter-Lorenz, Sylvester, Jonides, & Smith, 2003), involvement of this region most likely reflects the need to forgo a 'yes' response to a probe that is familiar in favor of a correct 'no' response. Finally, we also found increased bilateral IPS activity in response to lure items, consistent with McDermott et al.'s (2001) examination of this contrast in a standard DRM task. The role of PPC in SI and source memory is further discussed in our ROI analysis of this region.
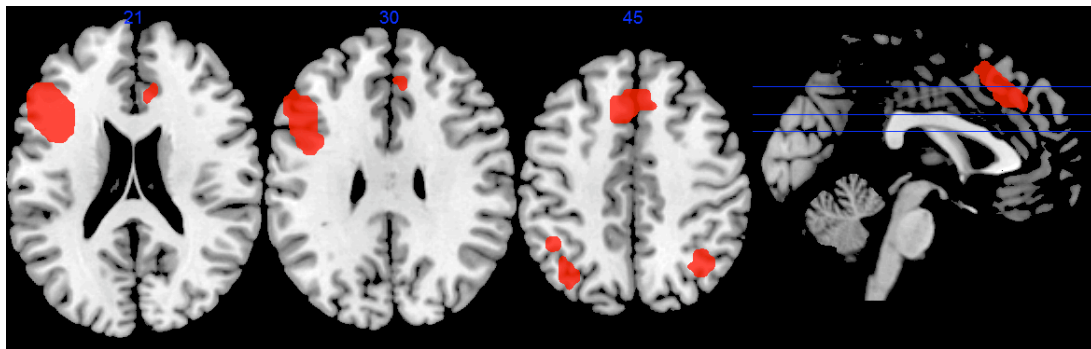
Figure 3.3.
Regions showing increased activity for correct rejection of lure vs. unrelated negative probes. As predicted, we found a large cluster of activation in L VLPFC associated with semantic interference. The peak of this activity, -48 21 21, was somewhat superior to those generally reported in tasks that examine proactive interference in the RP task (see text).

We examined neural mechanisms of true and false memory by first identifying regions associated with true and false recognition separately. For true recognition, we compared correct recognition of positive probes to correct rejections of unrelated negative probes (Pos. Hit > Neg. CR). For false recognition, we compared false alarms to lure probes to correct rejections of unrelated negative probes (Lure FA>Neg. CR).

True recognition was associated with increased activity a predictable network of fronto-parietal regions consistently associated with retrieval from verbal short-term

memory (Cohen, Perlstein, Braver, Nystrom, et al., 1997; D'Esposito, Postle, & Rypma, 1998; Rypma & D'Esposito, 1999; Chein & Fiez, 2001; Bedwell, Horner, Yamanaka, Li, et al., 2005). Most notably, these included large increases of activation in left anterior prefrontal/ dorsolateral prefrontal cortex (BA 10/46) and bilateral inferior parietal cortices (BA 40). False recognition showed a similar pattern of fronto-parietal activation, as well as a large cluster of activation in left posterior cingulate/ retrosplenial cortex, a region previously linked to phenomenological feelings of remembering that may be independent of retrieval accuracy (Wagner, Shannon, Kahn, & Buckner, 2005).

We examined regions that distinguished true from false recognition by directly contrasting activation associated with correct recognition of positive probes and false recognition of lures (Pos. Hit > Lure FA). Results are displayed in Figure 3.4. Compared with false recognition, true recognition was associated with a large cluster of increased activity extending from the left putamen to the anterior bank of the medial temporal lobe. We also found increased activation in right VLPFC, a region consistently implicated in inhibitory control across a variety of task contexts (Garavan, Ross, & Stein, 1999; Chikazoe, Konishi, Asari, Jimura, & Miyashita, 2007; Chikazoe, Jimura, Asari, Yamashita, et al., 2009), and in the left fusiform gyrus (see Table 3.2). Taken together, these findings suggest increased sensory processing, as well as the need to exert inhibitory control or increased task monitoring in order to support correct recognition of studied items in the face of interference.
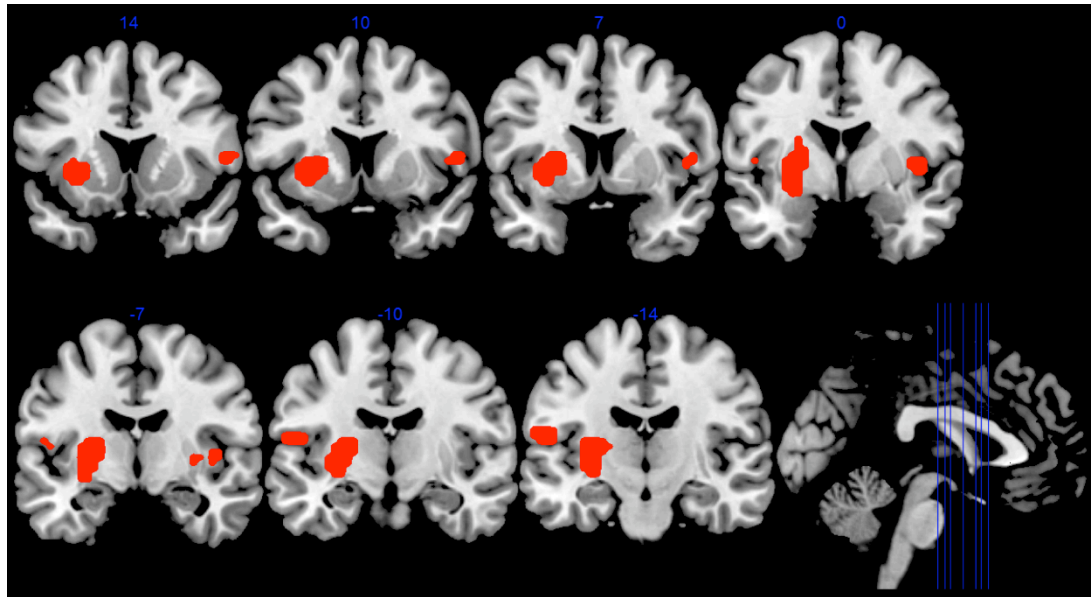
Figure 3.4.
Regions that dissociate true from false recognition (Pos. Hit>Lure FA). True recognition was associated with increased activity in the left putamen that extended to anterior MTL.

Region of interest (ROI) Analyses

Left mid-Ventrolateral Prefrontal Cortex

Results from our whole brain analysis supported our hypothesis that L VLPFC would contribute to resolving interference from related lures, although our activation peak was superior to the BA 45 peak linked to the resolution of PI. We conducted an ROI analysis to determine whether activation in the L VLPFC region associated with PI is also relevant for the control of SI. An ROI was formed by creating a 10mm sphere surrounding peak activation reported for the RN>NRN contrast in the RP task (MNI peak: -51 21 11; Jonides et al., 1998; Nelson et al., 2003).

Figure 3.5A plots mean percent signal change in L VLPFC for hits to positive probes, correct rejections of negative probes, and both correct rejections and false alarms to lure probes. Results show that activity in this ROI distinguished correctly rejected lures and unrelated negative probes, t=2.5, p<.05, *d*=.45. Furthermore, change in L VLPFC activity was strongly and positively correlated with individual differences in SI, r=.67, p<.01, (Figure 3.5B). This positive correlation between L VLPFC activity and SI is the same as that reported in the recent probes task (Jonides & Nee, 2006)[2].

We also found increased L VLPFC activation for lures that were falsely recognized, as compared to correctly rejected unrelated negative probes, suggesting that L VLPFC may be similarly recruited in response to SI, even when this interference is not successfully resolved. We can rule out the interpretation that L VLPFC is simply responding to the presence of semantic familiarity rather than interference per say, by examining activation associated with positive probes. Positive probes in our task are those that were studied in the context of a list of associated items, and should therefore be both semantically and episodically familiar, though this familiarity should not induce interference since it is consistent with a veridical recognition. If L VLPFC activity simply signaled the presence of familiarity, we would then expect increased activation in response to positive as well as lure probes. This is not the case, however, as probe-related activity did not differ for positive and unrelated negative probes in this region (p>.8).

---

[2] We also examined the correlation between individual variations in SI and changes in L VLPFC activity in a 10mm sphere surrounding our whole brain peak for the semantic interference contrast (-48 21 21). For this ROI, we found a similar, though slightly weaker, positive correlation between increased activity for lure vs. negative probes and increases in our behavioral SI measure, r=.51, p<.05.

Figure 3.5.
Mean percent signal change (PSC) in our L VLPFC ROI displayed as a function of trial type (A). Individual differences in L VLPFC activity for lure vs. unrelated negative probes was positively correlated with the RT index of semantic interference (B).

Right Posterior Parietal Cortex

Recent work has suggested an important role for posterior parietal regions in veridical memory processing and attention to task goals (Cabeza, et al., 2008; Ciaramelli, et al., 2008; Singh-Curry & Husain, 2009). Consistent with this assertion, our whole brain analysis revealed sites of increased activation in bilateral PPC associated with successful resolution of SI. In order to further examine the role of these parietal sites in SI, we examined brain-behavior relationships by correlating our behavioral measure of SI with activation changes in bilateral parietal peaks. ROIs consisted of 10mm spheres surrounding peak activations in right IPS (MNI peak: 38 -58 48) and left IPS (MNI peak: -34 -65 48). Results showed a trend toward a negative correlation between changes in right PPC activity and individual differences in SI, $r = .31$, $p < .1$, suggesting that subjects who showed greater increases in this ROI for lure vs. neg. probes showed less semantic interference. No such relationship was found for our left PPC region.

In order to further investigate a potential role for R PPC in interference resolution, we conducted an exploratory analysis to examine regions that positively correlated with activity in our R PPC ROI during successful resolution of SI. This was achieved by correlating activation changes in this region with changes in other regions throughout the brain on a voxel by voxel basis for each subject. Table 3.3 lists regions in which a minimum of 20 contiguous voxels were correlated at $r = .50$ or greater across all 20 subjects. High correlations between increases in right PPC activity and increases in bilateral fusiform and extrastriate regions are consistent with findings from our whole brain analysis of true vs. false memory, which showed large increases in fusiform activity associated with veridical recognition. High correlations were also found with medial

frontal regions (BA8) implicated in correct decision making in the face of uncertainty (Volz, Schubotz, & von Cramon, 2003, 2004; Huettel, Song, & McCarthy, 2005), and with left superior parietal cortex, a region that, together with right IPS has been linked to top-down control processes that allow for flexible management of task goals in the face of distraction (Ciaramelli, et al., 2008). Taken together, this correlational analysis supports the interpretation that right IPS is part of a network of regions that together work to support accurate task performance in the face of interference.

Table 3.3.
Regions correlated with increases in R PPC activity during successful resolution of SI .

	Peak			Voxels	r value	BA	Region
	x	y	z				
Occipital	41	-58	-15	361	0.64	37	Right Fusiform Gyrus
	-31	-72	-18	326	0.68	18/37	Left Fusiform Gyrus
	14	-28	-9	87	0.53		Right lingual gyrus
	-3	-93	3	65	0.56	18	Left extrastriate cortex
Frontal	-10	38	54	142	0.57	8	Left medial frontal lobe
	28	17	42	128	0.61	8	Right middle/inferior frontal gyrus
	21	65	9	56	0.6	10	Right superior frontal gyrus
Parietal	-21	-62	69	215	0.65	7	Left superior parietal lobe
Temporal	-55	-41	0	56	0.65	22	Left middle temporal gyrus
	-41	-34	15	36	0.65	41	Left superior temporal gyrus

Note- Regions listed contained a minimum of 20 contiguous voxels were correlated at r=.50 or higher across all 20 subjects.

Left frontal-polar cortex

An additional ROI analysis examined differences in a 10mm sphere surrounding the L FPC peak (MNI: -33 45 -9) reported by Badre and Wagner (2005). Activation in this region did not distinguish between any probe types in our task, and was not correlated with behavioral indices of SI or false memory. One possibility is that this

region, which Badre and Wagner (2005) attributed to episodic monitoring processes, is recruited for the resolution of interference due to temporal recency in the RP task, and may not be relevant to interference induced by items that have not been studied previously. Another possibility is that the number of trials utilized in the present investigation may not have provided us with adequate statistical power to defect probe-related differences in this region (see Badre & Wagner, 2005). Although we have no evidence for a L PFC contribution to SI, we cannot formally rule out this possibility.

DISCUSSION

The present study investigated the neural mechanisms of semantic interference and false recognition in a short-term variation of the DRM task (Deese, 1959; Roediger & McDermott, 1995; Atkins & Reuter-Lorenz, 2008). We examined the neural mechanisms of SI by identifying regions that showed increased activation for correct rejection of items related in meaning to current memoranda (lures) as compared to unrelated negative probes. In light of past research indicating a role for L VLPFC in proactive interference induced by recently studied items (Jonides & Nee, 2006), we predicted this region would also be involved in interference induced by the lure items, which were never studied but which were semantically familiar. Consistent with this prediction, we found increased L VLPFC (BA45) activity associated with the correct rejection of lure vs. unrelated negative probes. ROI analyses revealed a strong positive correlation between individual differences in the magnitude of SI and probe-related increases in L VLPFC (see Figure 3.5). This finding demonstrates that the positive relationship between behavioral indices of interference and L VLPFC activity is not

unique to temporal familiarity, and suggests that common neural substrates are engaged in response to interference induced by either temporal or semantic familiarity.

Positive correlations between behavioral measures of interference and increased activity in L VLPFC could be interpreted as either a) evidence this region is the site of interference resolution and therefore must work harder to resolve interference as it increases, or b) this region creates an index of interference in the environment that is used by other cortical regions in the service of accurate memory retrieval. We attempted to distinguish between these possibilities by comparing L VLPFC activity for lures that were ultimately rejected to those that were falsely recognized. Results showed a similar increase in probe-related activity for both correct rejections and false alarms to lures, indicating that the L VLPFC responded equivalently to the presence of interference that was ultimately resolved successfully, and interference that was not resolved.

False alarm rates in RP tasks used to investigate the neural mechanisms of PI are normally quite low. As such, these investigations have focused almost exclusively on the successful resolution of interference, and have generally interpreted L VLPFC increases in this context. Our findings are consistent with the interpretation that L VLPFC responds to the selection demands associated with multiple active representations in memory (Thompson-Schill, et al., 1997; Badre & Wagner, 2007), but suggest that the region responds equivalently when this selection process is unsuccessful.

In comparison to L VLPFC, our exploratory analysis of brain-behavior correlations in R PPC showed a marginal negative correlation between changes in activation for lure vs. negative probes and reductions in RT measures of semantic interference. This region, located along the superior bank of the IPS, has also been

implicated in correct rejections of lure vs. unrelated negative probes in an episodic DRM task (McDermott, et al., 2000), and in source memory across a variety of tasks (Cabeza, et al., 2008; Ciaramelli, et al., 2008; Mitchell & Johnson, 2009). Furthermore, results from our correlational analysis show that increases in R PPC is correlated with increases in activity in a number of regions shown to support accurate responding in the face of uncertainty and distraction, including bilateral fusiform, medial frontal and left superior parietal regions (Table 3.3).

Our interpretation of R PPC's role in the current task is consistent with recent work suggesting distinct roles for IPS/superior parietal regions, associated with top-down attention to memory, and inferior parietal regions, associated with bottom-up attentional capture by retrieved memory representations (Ciaramelli, et al., 2008). The positive correlation between activation in our PPC ROI, which lies along the superior boarder of the IPS, and decreased vulnerability to false recognition, is consistent with the hypothesis that this region mediates top-down control of access to memory representations that may help prevent attentional capture by unstudied, but familiar, lures.

Taken together, our results suggest a possible distinction between roles for L VLPFC and R PPC in the resolution of semantic interference. L VLPFC shows increased activation in response to semantic lures, but does not distinguish between lures that are ultimately correctly rejected and those that are falsely recognized as studied items. Our finding of negative correlations between R PPC activity and interference suggests this region may play a central role in adjudicating the familiarity of lure items in the service of timely and accurate task performance.

ACKNOWLEDGEMENTS

REFERENCES

Atkins, A. S., & Reuter-Lorenz, P. A. (2005). *'False' working memories: Memory distortion in a mere 4 seconds.* Paper presented at the Psychonomic Society 46th Annual Meeting, Toronto, ON.

Atkins, A. S., & Reuter-Lorenz, P. A. (2008). False working memories? Semantic distortion in a mere 4 seconds. *Memory & Cognition, 36*(1), 74-81.

Badre, D., & Wagner, A. D. (2005). Frontal Lobe Mechanisms that Resolve Proactive Interference. *Cereb. Cortex, 15*(12), 2003-2012.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901.

Bedwell, J. S., Horner, M. D., Yamanaka, K., Li, X. B., Myrick, H., Nahas, Z., et al. (2005). Functional neuroanatomy of subcomponent cognitive processes involved in verbal working memory. *International Journal of Neuroscience, 115*(7), 1017-1032.

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: an attentional account. *Nature Reviews Neuroscience, 9*(8), 613-625.

Chein, J. M., & Fiez, J. A. (2001). Dissociation of verbal working memory system components using a delayed serial recall task. *Cerebral Cortex, 11*(11), 1003-1014.

Chikazoe, J., Jimura, K., Asari, T., Yamashita, K., Morimoto, H., Hirose, S., et al. (2009). Functional Dissociation in Right Inferior Frontal Cortex during Performance of Go/No-Go Task. *Cerebral Cortex, 19*(1), 146-152.

Chikazoe, J., Konishi, S., Asari, T., Jimura, K., & Miyashita, Y. (2007). Activation of right inferior frontal gyrus during response inhibition across response modalities. *J Cogn Neurosci, 19*(1), 69-80.

Ciaramelli, E., Grady, C. L., & Moscovitch, M. (2008). Top-down and bottom-up attention to memory: A hypothesis (AtoM) on the role of the posterior parietal cortex in memory retrieval. *Neuropsychologia, 46*(7), 1828-1851.

Coane, J. H., McBride, D. M., Raulerson, B. A., & Jordan, J. S. (2007). False memory in a short-term memory task. *Experimental Psychology, 54*(1), 62-70.

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature, 386*(6625), 604-608.

D'Esposito, M., Postle, B. R., & Rypma, B. (1998, Dec 01-03). *Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies.* Paper presented at the Workshop on Executive Control and the Frontal Lobe, Delmenhorst, Germany.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17-22.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved Assesment of significant activation in functional magnetic resonance imaging (fMRI) - use of a cluster size threshold. *Magnetic Resonance in Medicine, 33*(5), 636-647.

Garavan, H., Ross, T. J., & Stein, E. A. (1999). Right hemispheric dominance of inhibitory control: An event-related functional MRI study. *Proceedings of the National Academy of Sciences of the United States of America, 96*(14), 8301-8306.

Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage, 9*(4), 416-429.

Gold, B. T., & Buckner, R. L. (2002). Common Prefrontal Regions Coactivate with Dissociable Posterior Regions during Controlled Semantic and Phonological Tasks. *Neuron, 35*(4), 803-812.

Hancock, T., & Hicks, J. (2002). Backward associative strength determines source attributions given to false memories. *Psychonomic Bulletin & Review, 9*(4), 807-815.

Hirshorn, E. A., & Thompson-Schill, S. L. (2004, Jun 13-17). *Role of the left inferior frontal gyrus in covert word retrieval: Neural correlates of switching during verbal fluency.* Paper presented at the 10th Annual Meeting on Human Brain Mapping, Budapest, HUNGARY.

Huettel, S. A., Song, A. W., & McCarthy, G. (2005). Decisions under uncertainty: Probabilistic context influences activation of prefrontal and parietal cortices. *Journal of Neuroscience, 25*(13), 3304-3311.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139*(1), 181-193.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences of the United States of America, 95*(14), 8410-8413.

McDermott, K. B., Jones, T. C., Petersen, S. E., Lageman, S. K., & Roediger, H. L. (2000). Retrieval success is accompanied by enhanced activation in anterior prefrontal cortex during recognition memory: An event-related fMRI study. *Journal of Cognitive Neuroscience, 12*(6), 965-976.

Mitchell, K. J., & Johnson, M. K. (2009). Source Monitoring 15 Years Later: What Have We Learned From fMRI About the Neural Mechanisms of Source Memory? *Psychological Bulletin, 135*(4), 638-677.

Monsell, S. (1978). Recency, Immediate Recognition Memory, and Reaction-Time. *Cognitive psychology, 10*(4), 465-501.

Nelson, J. K., Reuter-Lorenz, P. A., Persson, J., Sylvester, C. Y., & Jonides, J. (2009). Mapping interference resolution across task domains: a shared control process in left inferior frontal gyrus. *Brain Res, 1256*, 92-100.

Nelson, J. K., Reuter-Lorenz, P. A., Sylvester, C. Y., Jonides, J., & Smith, E. E. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proc Natl Acad Sci U S A, 100*(19), 11171-11175.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional Specialization for Semantic and Phonological Processing in the Left Inferior Prefrontal Cortex. *NeuroImage, 10*(1), 15-35.

Roediger, H. L., & McDermott, K. B. (1995). Creating false Memories: remembering words not presented in lIsts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803.

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

Rypma, B., & D'Esposito, M. (1999). The roles of prefrontal brain regions in components of working memory: Effects of memory load and individual differences. *Proceedings of the National Academy of Sciences of the United States of America, 96*(11), 6558-6563.

Singh-Curry, V., & Husain, M. (2009). The functional role of the inferior parietal lobe in the dorsal and ventral stream dichotomy. *Neuropsychologia, 47*(6), 1434-1448.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*(26), 14792-14797.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2003). Predicting events of varying probability: uncertainty investigated by fMRI. *Neuroimage, 19*(2), 271-280.

Volz, K. G., Schubotz, R. I., & von Cramon, D. Y. (2004, May 25-27). *Variants of uncertainty in decision-making and their neural correlates.* Paper presented at the 2nd Conference on NeuroEconomics, Munster, GERMANY.

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences, 9*(9), 445-453.

Chapter IV

RESOLVING SEMANTIC AND PROACTIVE INTERFERENCE IN SHORT-TERM MEMORY

Introduction

Interference is one of the greatest burdens on the cognitive system. It's why we put the cereal box in the refrigerator when we have an idle thought, but it is also related to more serious lapses of attention, which can lead to significant accidents. Additionally, interference has profound adverse affects on short-term memory, and is one of the major sources of forgetting in short-term memory (Nairne, 2002; Wixted, 2005; Lewandowsky, Geiger, & Oberauer, 2008; Berman, Jonides, & Lewis, 2009).

To make matters worse, short-term memory is susceptible to multiple forms of interference. Classical investigations demonstrated acoustic and visual similarity interference effects both on short-term memory recall and recognition, wherein acoustically or visually similar items were incorrectly substituted for memorized items at retrieval (Conrad, 1964; Laughery & Harris, 1970; Logie, Della Sala, Wynn, & Baddeley, 2000). Early work also established that the semantic context of memoranda lead to a build-up of interference. In classic release from proactive interference (PI) paradigms (Wickens, 1970), recall and recognition of items on a current memory set declines as a

function of the number previous lists studied from the same category. When the category is changed, PI is eliminated (Kincaid & Wickens, 1970; Wickens, 1970).

More recent investigations of PI have highlighted the strong influence of episodic familiarity on short-term memory retrieval. One task used to test the influence of proactive interference is the recent probes (RP) task. In the task (Monsell, 1978; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998), participants study a set of items that are retained for several seconds. After the retention interval, subjects are presented with a probe to which they must make a Y/N response. Critically, negative probes are of two kinds: recent negative (RN) and non-recent negative (NRN). RN probes are items that are not present on the current trial, but were present in the memory set on the trial immediately preceding the current one, while NRN items are probes that have not appeared on the last several trials. Investigations of the RP task consistently show that participants are markedly slower and more error-prone when rejecting RN relative to NRN probes. These findings have been interpreted as evidence for the need to engage in interference resolution, a controlled cognitive process by which the episodic familiarity of RN probes is adjudicated in service of veridical retrieval from memory (Jonides & Nee, 2006).

Other recent work has demonstrated semantic interference (SI) in short-term memory, wherein correct rejection of unstudied items related in meaning to those in memory is slower and more error-prone than rejection of unrelated items (Atkins & Reuter-Lorenz, 2008). In a recent investigation, we used a short-term variant of the Deese-Roediger-McDermott paradigm (DRM; Roediger & McDermott, 1995) to examine false memory and semantic interference (Atkins & Reuter-Lorenz, 2008). In this ST-

DRM task, participants studied sets of 4 words which were all associates of a common unstudied theme word. Following a 3-4 second retention interval, which was either unfilled or filled with a math verification task (i.e., a distraction task), participants saw a single probe word to which they made a Y/N response to indicate whether the probe was a member of the current memory set. Critically, negative probes were of two kinds: unrelated negative (NEG) probes which were not members of the current set and which were unrelated in meaning to the memorized items, and related negative (LURE) probes which consisted of the unstudied theme associated with items in the current set. Each trial consisted of a memory set associated with a trial-unique theme, with no repetition of themes or memoranda across trials. Findings showed reliable false recognition of lure probes, and high levels of interference associated with correct rejection of these items relative to unrelated negative probes. This suggests that like temporal recency, semantic context can induce interference in short-term memory.

The RP and ST-DRM paradigms used to investigate PI and SI respectively, both require participants to negotiate the familiarity of critical probe items (RN or LURE probes) in service of accurate memory retrieval. In the first case, RN probes are activated due to their recent status as memoranda, while in the second, LURE probes become active at encoding via spreading activation through a semantic network (Underwood, 1965; Collins & Loftus, 1975; Roediger, McDermott, & Robinson, 1998). One possibility is that both the episodic and semantic context of critical probes may promote source memory confusion, increasing the likelihood that such probes will be misattributed to the current memory set (Nelson, Reuter-Lorenz, Sylvester, Jonides, & Smith, 2003). This account is consistent with feature-based models of memory such as

that proposed by Nairne (2002). According to this model, successful memory retrieval relies on accurate recombination of activated cues associated with various features of memoranda, such as episodic, semantic, lexical and phonological features. A reduction in the distinctiveness of episodic or semantic features that distinguish probes from memoranda will induce interference. Under such circumstances, a similar interference resolution process may be required. When such a process fails, RN or LURE probes will be falsely recognized as members of the current memory set; when it succeeds, correct rejection of critical probes is slowed due to the need to engage in interference resolution.

Mechanisms of interference resolution in the RP task have been a subject of numerous neuroimaging investigations that have highlighted a role for left ventrolateral prefrontal cortex (L VLPFC) in this and other cognitive control tasks (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Jonides, et al., 1998; Nelson, et al., 2003; Badre & Wagner, 2005; Jonides & Nee, 2006; Nee, Wager, & Jonides, 2007). In contrast, there is a paucity of work that examines the resolution of interference from multiple sources (though see Nee, Jonides, & Berman, 2007; Oztekin, Curtis, & McElree, 2008; Nelson, Reuter-Lorenz, Persson, Sylvester, & Jonides, 2009), and to our knowledge, no previous investigation has directly examined the influence of semantic context on the recent probes measure of PI in short-term memory.

The purpose of the present work, therefore, was to systematically examine the relationship between PI and SI in short-term memory, and test the hypothesis that a common resolution process may be recruited to resolve both types of interference. Exp. 1 examines the effects of articulatory suppression (AS) during the retention interval of

short-term memory tasks designed to assess each form of interference. The effect of AS is of interest for two reasons, described below.

First, previous investigations of PI have shown this effect to be insensitive to a variety of manipulations, including the requirement to engage in AS during the inter-trial interval (ITI) of the RP task (Berman, Jonides & Lewis. 2009). Lengthening the ITI also did not alter the PI effect appreciably, nor did instructing participants to attempt to ignore past sets (Berman, Jonides & Lewis, 2009). These results show that participants are not rehearsing items from the past set, and that interference from previously studied items is not readily mitigated by strategic executive control. We predicted that AS in the retention interval would increase the PI effect by decreasing the signal to noise ratio between trials on the current and past memory sets, thus making participants less able to distinguish RN probes from current memoranda. If so, participants should show increased PI on AS trials, as well as an increased tendency to incorrectly recognize RN probes as items on the current set.

Second, our previous investigations of SI have shown this measure to be insensitive to the requirement to engage in a math verification task during the retention interval (Atkins & Reuter-Lorenz, 2008). We predicted the SI would in fact increase if our distraction task more directly interfered with subvocal rehearsal of current memory set items. If interference elicited by lure probes in our ST-DRM task is resolved by the same process as PI, then AS during the retention interval should also increase SI by the same logic that predicts an increase in PI. In both cases, AS should interfere with verbatim rehearsal processes (Levy, 1971; Baddeley, Thomson, & Buchanan, 1975),

reducing the distinctiveness of current memoranda by decreasing the signal to noise ratio of these items relative to recent negative or lure probes.

The results from Exp. 1 indicate that AS increases both proactive and semantic interference effects. These findings are consistent with our predictions that interfering with the rehearsal of items on the present memory set increases interference elicited by recent negative probes, which were previously studied but irrelevant to the current trial, and from lure probes, which were never studied but have become active through spreading semantic activations.

In Exps. 2 and 3, we directly examine the relationship between SI and vulnerability to PI in a variant of the recent probes task. If the semantic relationship between the memory set and the probe can induce interference, perhaps semantic contextual tags can be leveraged to reduce PI in cases where the semantic context of a RN probe is incongruent with that of the current memory set. By the same logic, PI may increase when a RN probe shares a semantic context with current memoranda.

Consistent with these predictions, we find PI is reliably reduced, though not eliminated, when recent negatives are drawn from a separate semantic category than memoranda on the current trial, suggesting only limited release from proactive interference associated with the category shift. Additionally, PI reliably increased when the semantic categories matched, suggesting that semantic context can change the degree to which we are vulnerable to proactive interference. Within the additive factors framework (Sternberg, 1969), this interaction between semantic and episodic familiarity supports the notion that common cognitive process may be recruited to resolve interference from both sources.

EXPERIMENT 1

Method

Participants

Thirty-two participants (11 males; mean age=19.6 years) were recruited from the University of Michigan. All participants gave informed consent as reviewed by the university's Institutional Review Board. Participants were paid $15 per hour for their participation plus bonuses for fast and accurate responding throughout the experiment. Bonus scores were calculated on a trial-by-trial basis with the following equation:

$$\text{Trial Score} = \text{Probe\_ACC} * (700\text{ms} - \text{Probe\_RT})$$

where probe accuracy (ACC) is a binary variable, 1 if correct and 0 if incorrect, and RT is response time. Individual trial scores were summed together to yield a total score. Participants were paid a penny for each point of their total score.

Design/Procedure

In order to examine the effects of AS on proactive and semantic interference, we utilized a 2X2 mixed design, with interference type (PI vs. SI) as a between subjects variable and AS as a within subjects variable. As such, half of all participants (n=16) completed the recent probes task (to assess PI), and half (n=16) completed the ST-DRM task (to assess SI). Participants in each group completed two task blocks, one with AS and one without. The order of Articulatory Suppression (AS) and No Suppression (no-AS) blocks was counter-balanced across subjects.

Recent probes task

We used a recent-probes (RP) task to assess the effects of proactive interference (Fig. 1A). On each trial a set of 4 words was displayed for 2000ms. Following a variable length 3000-4000ms retention interval, a single probe word was displayed. Participants needed to respond affirmatively if the probe was one of the words on the set, and negatively if not. The 'No' trials were of interest and came in two forms. The first type was a *recent negative* (RN) trial, in which the probe was not a member of the current set, but was a member of the previous set. The second type was a standard *nonrecent negative* (NRN) trial, in which the probe was neither a member of the current set, nor a member of the two previous sets. Previous work using this task has shown that participants are both delayed and less accurate when responding to these RN compared to NRN probes, and this difference in performance is an assay of proactive interference. We also utilized two forms of positive probes: *recent positives* (RPOS), which were present on the current memory and immediately preceding memory set, and *non-recent positives* (NRPOS), which were present in the current set, but not present in the two preceding trials.

Short-term DRM task

We used a short-term DRM paradigm (ST-DRM; Atkins & Reuter-Lorenz, 2008) to examine semantic interference (Fig. 1B). The progression of each trial was similar to that of the RP task, except that memory sets consisted of 4 semantically related items, all associated with a common theme word. As in the RP task, participants made a Y/N judgment in response to a probe presented following the retention interval. During this

task, theme words served as the probes on all the trials. As in the RP task, there were two variations of "No" trials. The first were *unrelated negative* (NEG) trials, in which the probe consisted of a theme word associated with a nonpresented list. The second were *lure* (LURE) trials, in which the probe consisted of the (unstudied) theme associated with the present memory set. On *positive* (POS) trials, the associated theme was embedded in the memory set, and served as the positive probe. No participant was exposed to a given theme or probe more than once during the experiment. The mean backward associative strength (BAS), a normed measure of the semantic association between each memoranda and the theme word (see Roediger, Watson, McDermott, & Gallo, 2001; Hancock & Hicks, 2002), was equated across memory lists. Probe type was counterbalanced with lists across participants. This procedure ensured that participants encountered the same probes (all theme words) but in different contexts: as lures, unrelated negative probes, or positive probes. Trials were presented in random order for each participant.

Articulatory Suppression

On AS blocks, participants counted aloud from 1 to 3 repeatedly at a rate of approximately four utterances per second. Counting began at the onset of a blue fixation cross which appeared just after encoding and stayed on the screen throughout the retention interval. The pace of counting was practiced prior to completion of the AS trial block. During the experiment, counting was monitored to ensure compliance with task instructions. AS and non-AS blocks were administered in a counterbalanced order across participants.
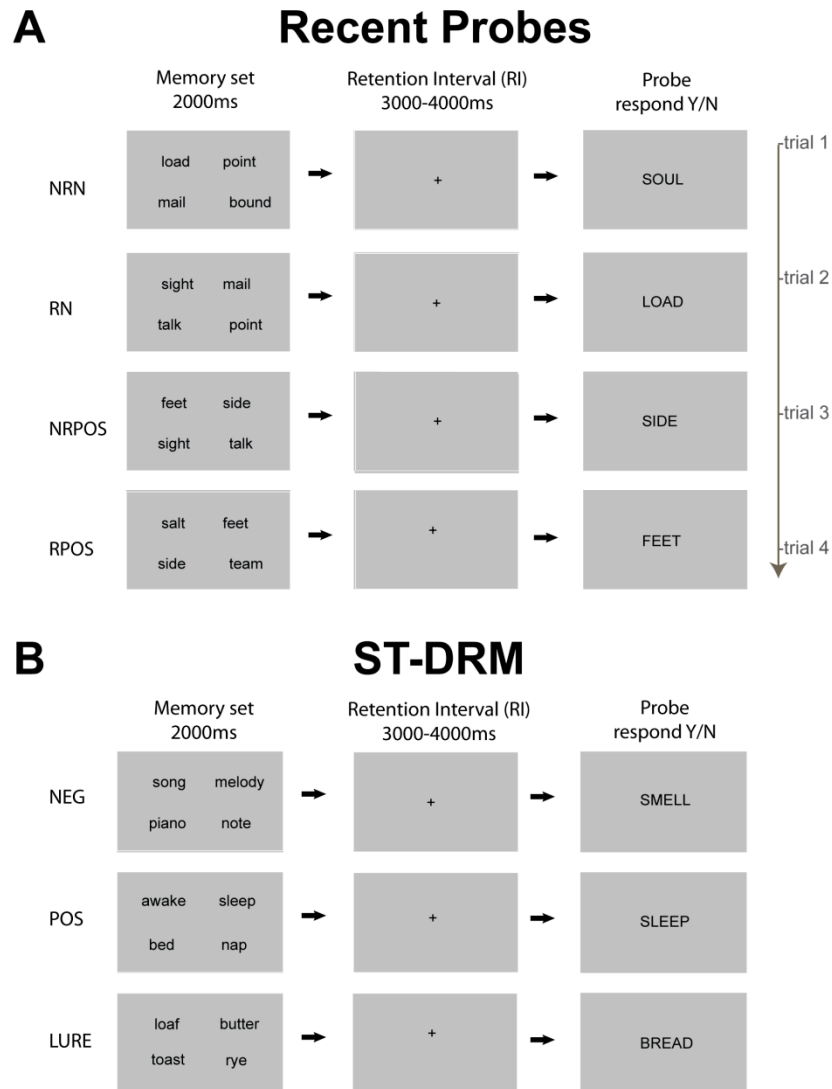
Figure 4.1.
The Recent Probes and ST-DRM tasks used in Exp. 1. In the RP task (A) recognition probes could be either non-recent negatives (NRN), recent negatives (RN), non-recent postives (NRPOS), or non-rencent postives (NRPOS). In the ST-DRM task (B), probes could be either negative (NEG), postive (POS), or negative lure (LURE). During AS blocks for both tasks, the fixation point presented during the RI was bright blue rather than black. Participants began counting aloud from 1 to 3 repeatedly at the onset of this symbol and continued throughout the retention interval. See text for details.

Results and Discussion

The results from Exp. 1 were consistent with our initial predictions. Both PI and SI effects significantly increased with the requirement to engage in AS during the retention interval. Mean accuracy and median response times (RTs) for correct responses to each probe type in each task are presented in Figure 4.2. In order to minimize the influence of outliers in our RT analyses, median RTs for each probe type were computed for each subject; mean group RT measures were computed from these medians.

Proactive interference

In order to examine the effects of probe recency and AS in the RP task, 2X2X2 mixed effects analyses of variance (ANOVAs) were conducted separately for positive and negative probes. Within subjects variables were Probe Recency (recent vs. non-recent) and AS condition (AS vs. no-AS). Task Order (AS first vs. AS second) was included as a between subjects variable. With respect to negative probes, we found main effects of Probe Recency on both accuracy, $F=33.11$, $p<.001$, $\eta2=.70$, and RT, $F=53.19$, $p<.001$, $\eta2=.79$, indicating that participants were both slower and less accurate in rejecting RN relative to NRN probes. These findings replicate others that have demonstrated PI using similar paradigms, and indicate that our task succeeded in inducing PI. Critically, we found a significant Probe Recency by AS interaction with respect to both accuracy, $F=13.86$, $p<.005$, $\eta2=.50$, and RT, $F=42.49$, $p<.001$, $\eta2=.75$, indicating that PI effects on both dependent measures increased with the requirement to engage in AS. Our RT index of PI (RN-NRN) increased from a mean of 61.34 ms in the

non-AS condition to a mean of 134.25 ms in the AS condition. Our accuracy measure of PI (NRN-RN) similarly increased from .04 in the non-AS condition to .15 in the AS condition.
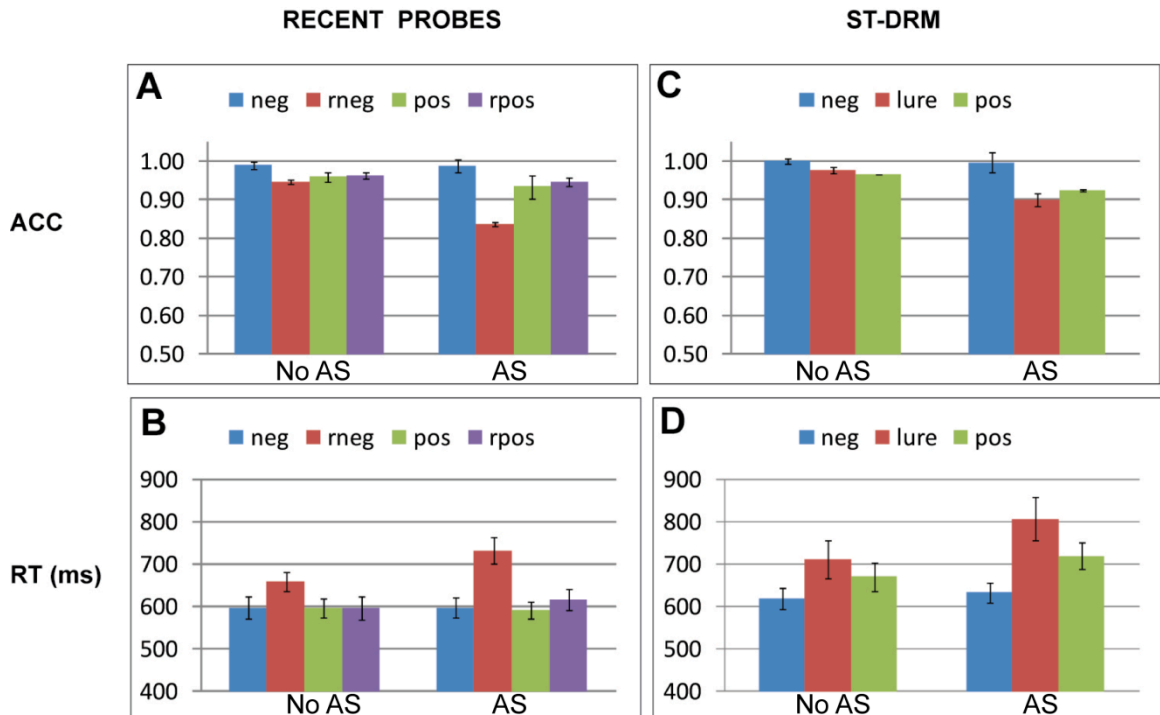

Figure 4.2.
Mean accuracy and median RTs for the Recent Probes (Panels A and B) and ST-DRM tasks (Panels C and D) used in Exp.1. Articulatory supression (AS) reliably increased both accuracy and RT indices of PI and SI.

Collapsing across recent and non-recent negative probes, participants we less accurate, $F=13.27$, $p<.005$, $\eta2=.49$, but not reliably slower ($p>.1$) on AS blocks. Similarly, with respect to positive probes, participants were less accurate, $F=8.65$, $p<.05$, $\eta2=.38$, but not slower ($p>.1$) in responding to positive probes in AS blocks.

Examination of RTs (Figure 4.2B) shows that AS selectively increased the time required to correctly reject RN probes. Post-hoc paired t-tests showed no reliable increase in RT for any other probe type. This indicates that AS selectively affected

interference induced by RN probes, and was not associated with a general slowing of responses to other probe types.

Semantic interference

Mean accuracy and response times (RTs) for the ST-DRM task are presented in Figure 4.2, panels C and D, respectively. Repeated measures ANOVAs were conducted to assess the effects of negative probe type (neg. vs. lure), AS condition, and Task Order on both dependant measures. Participants were less accurate in responses to lures, relative to negative unrelated probes, $F = 25.20$, $p<.001$, $\eta2=.64$. Furthermore, correct rejections of lure probes were slowed relative to negative unrelated probes, $F=29.01$, $p<.001$, $\eta2=.67$. These main effects replicate our previous work (Atkins & Reuter-Lorenz, 2008) and demonstrate the effectiveness of the ST-DRM task in producing reliable measures of SI.

Main effects of AS were found for negative probe accuracy, $F=6.26$, $P<.05$, $\eta2=.31$, and RT, $F=10.70$, $p<.05$, $\eta2=.31$, indicating that, collapsing across lure and unrelated negative probes, participants were both less accurate and slower to correctly respond in the AS block. Critically, we found reliable Probe Type by AS interactions on both probe accuracy, $F= 5.81$, $p<.05$, $\eta2=.29$, and RT, $F=8.42$, $p<.05$, $\eta2=.34$, indicating that both accuracy and RT measures of SI increased with AS.   Our RT index of SI (LURE-NEG) increased from a mean of 92 ms in the no-AS condition to a mean of 173.88 ms in the AS condition. Our accuracy measure of SI (NEG-LURE) also increased from .02 in the no-AS condition to .10 in the AS condition.

Taken together, results from Exp. 1 show that AS increases both proactive and semantic interference effects. One possibility is that articulatory suppression may reduce the availability of implicitly generated contextual tags that could be used to discern whether or not a probes' familiarity is due to its membership in the current memory set (Nairne, 2002). Such contextual tags need only to be consulted when the item-specific and familiarity-based memory processes conflict, as they do on recent probe trials. Another possibility is that AS may reduce the distinctiveness of current memoranda, lowering the signal to noise ratio between these items and those active due to previous study or spreading semantic activation. Positive probes may maintain sufficient activation to support timely recognition, but this reduction in distinctiveness may lead to increased interference from RN and LURE probes.

These findings are consistent with our predictions and suggest that interrupting rehearsal of items in the present memory set increases interference: 1) elicited by RN probes, which were previously studied but irrelevant to the current trial, and 2) from lure probes, which were never studied but have become active through spreading semantic activations. This similar influence of the AS manipulation on both forms of interference suggests that a common cognitive control mechanism may be used to resolve interference from familiar probes, regardless of whether they are familiar because of recent study (temporal or episodic familiarity) or because they share a semantic context with items currently held in memory.

EXPERIMENT 2

Results from Exp. 1 indicate that the requirement to engage in articulatory suppression during the retention interval increases both proactive and semantic interference in short-term memory. Although the similar effects of AS on both PI and SI are consistent with the notion that similar psychological processes may underlie the resolution of both forms of interference, findings from Exp. 1 cannot speak directly to this possibility.

In order to examine more directly the relationship between proactive and semantic interference, we conducted two additional experiments. The presence of SI in short-term memory suggests that the semantic context of probes items is used in the recognition decision process. Lure probes take longer to reject and are more likely to be falsely recognized because they are related in meaning to items currently in memory. If the semantic relationship between the memory set and the probe can induce interference, perhaps semantic contextual tags can be leveraged to reduce PI in cases where the semantic context of a recent negative probe is incongruent with that of the current memory set.

In Exp. 2 we explored this possibility in a paradigm that incorporates semantic categories into a standard RP task. Participants completed two versions of the RP task, one of which utilized memory sets in which all memoranda were exemplars of a single semantic category, and one version in which they were not (i.e., the standard RP task). If semantic information can be used to reduce PI, we would expect a reduction in PI when

RN probes are drawn from a different semantic category than items in the current memory set. [3]

Method

Participants

Ten participants (5 males; mean age 20.7 years) were recruited from the University of Michigan to participate in this study. All participants gave informed consent as reviewed by the University of Michigan's Institutional Review Board. As in Exp. 1, participants were paid $15 per hour for their participation plus bonuses for fast and accurate responding throughout the experiment.

Design/Procedure

We used a modification of the recent-probes task (Exp.1) to jointly assess the effects of semantic and proactive interference within the same paradigm. In order to incorporate semantic familiarity into the recent probes task, memory sets were drawn from two semantic categories, fruits and countries. There were 50 words utilized as memoranda: 25 fruit words and 25 country words. In the *categorized* (CAT) version of the task (Figure 4.3A), memory sets consisted of either all fruit words or all country words, with the category alternating from trial to trial in a predictable fashion (e.g., country-fruit-county…). Because categories alternated on each trial, recent negative probes in this task were always members of a different category than the current memory

---

[3] Exp. 2 was initially conceived by my collaborator Marc G. Berman, who was interested in the examining the standard recent probes effect with categorized stimuli. The experiment is included in this chapter because of its relevance to Exps. 1 and 3 which were motivated by my interest in understanding the relationship between PI and SI. As discussed subsequently, Exp.3 extends and replicates Exp. 2, and was conceived collaboratively with Berman, but motivated more directly by my effort to analyze the relationship between PI and SI.

set. Non-recent negative probes, on the other hand, could be either category-congruent or category-incongruent with memoranda. In the *mixed* (MIX) version of the task (Figure 4.3B), we used the same words, but memory sets contained a mixture of fruit and country names, making categorization of any single memory set impossible. In this task probes were drawn equally from both semantic categories.

There were 192 trials in both the Mixed and Categorized versions of the task. In the Mixed version, there were 96 POS trials, 48 RN trials and 48 NRN trials. For the sake of clarity, we will refer to these as POS_mixed, RN_mixed, NRN_mixed and NRN_mixed trials, respectively. In the Categorized version of the experiment there were 96 POS_match trials (that were items on the current memory set), 48 RN_mismatch trials (that were always from a different semantic category from the current memory set), 24 NRN_match trials (that were from the current memory set's semantic category) and 24 NRN_mismatch trials (that were from a different semantic category from the current memory set). In these two versions of the task there were no Recent Positive trials.

Exp. 2 utilized a within subjects design. All participants completed both versions of the task. The order of the task was counter-balanced across participants. This procedure allowed us to compare PI measures across two contexts, one in which semantic information could feasibly be used to improve task performance (mismatch trials) and one in which it could not (mixed trials). If semantic category information can be leveraged to reduce PI when RN probes are from a different semantic category than memory set items, then we should find a reduction in PI on mismatch trials. More specifically, RN_mismatch trials should be faster and less error prone than RN_mixed trials. Our design also allowed us to explore SI effects on trials in which NRN probes

were not members of the current set, but were members of the same semantic category.

(NRN_match trials should be more difficult than NRN_mismatch trials because there is

semantic interference). This measure of SI may be less robust than that observed in the

ST-DRM task in which memory sets were constructed from direct associates to lure

items. However, previous work on the associative nature of semantic memory has

demonstrated semantic priming using category exemplars (Collins & Quillian, 1969;

Neely, 1976; Huttenlocher & Kubicek, 1983). We therefore predicted that the categorical

relationship between memory items would be sufficient to induce spreading activation to

unstudied category exemplars, thereby increasing their familiarity.
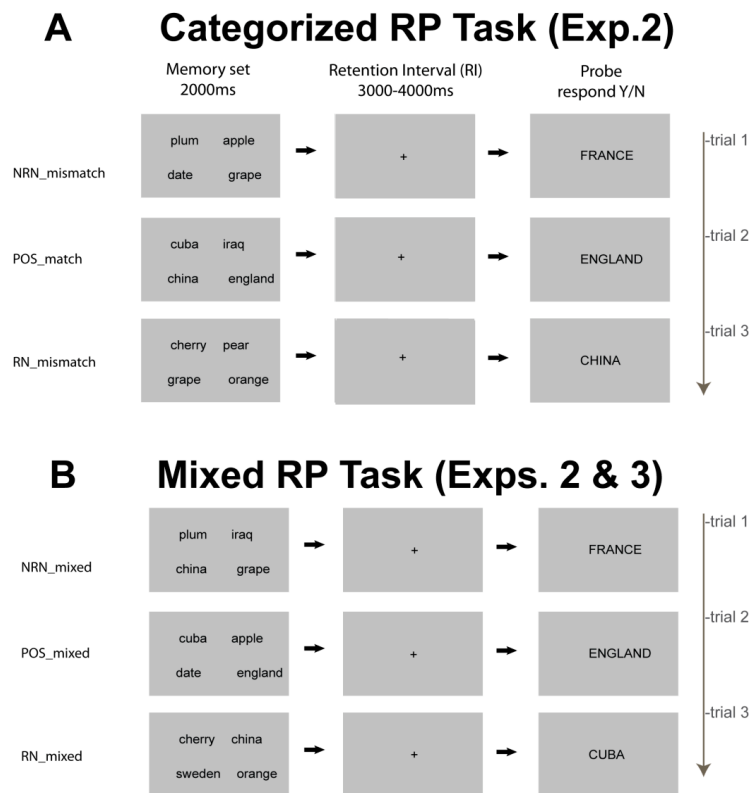
Figure 4.3.
The Categorized RP task used in Exp. 2, and Mixed RP task used in both Exp. 2 and Exp. 3.
In both versions of the task, memoranda consisted of fruit and country words. In the CAT
task (A) each memory set consisted of either all fruit or all country words. The category of
memory sets alternated predictably from trial to trial (fruit-country-fruit-country). In MIX
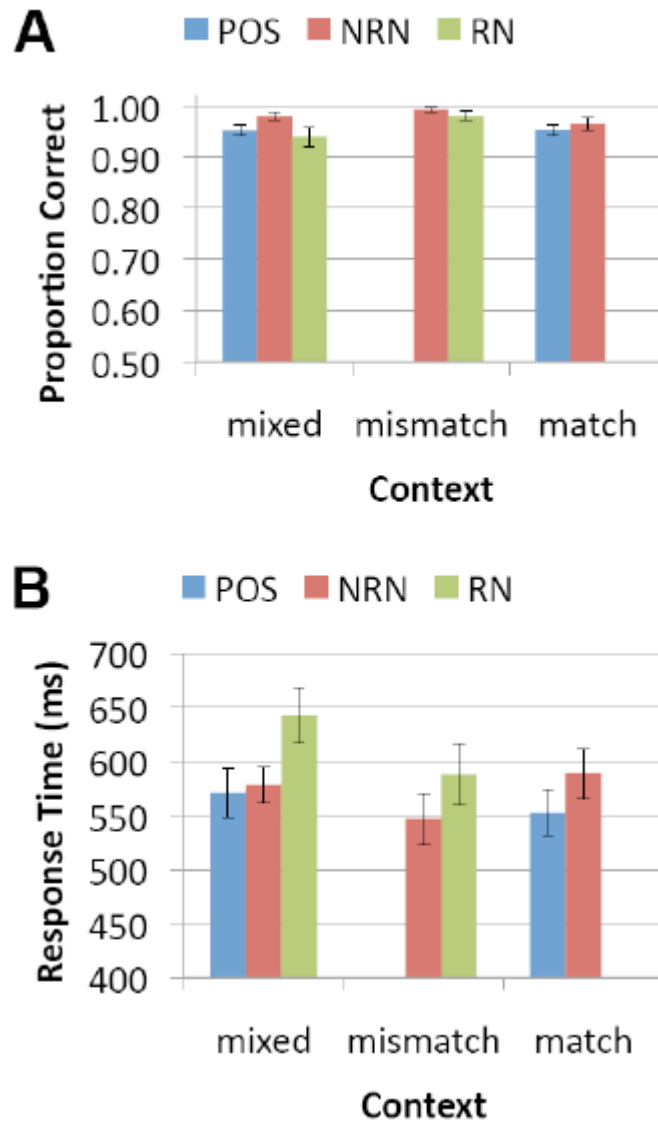task (B) memory sets contained a mixture of words from both categories.

Figure 4.4.
Mean accuracy and median RTs from Exp. 2. On mixed trials, which were
blocked, memory sets included a mixture of words from two semantic categories (fruits and
countries), and recent negative probes were drawn from each category with equal probability.
On both mismatch and match trials (which appeared in the categorized block), memory sets
contained items from only one semantic category. In Exp. 2, the semantic category of
memory sets in the categorized block alternated from trial to trial. As such, recent negative
(RN) probes were always drawn from the semantic category that was not presented on the
current trial (mismatch). Non-recent negative trials were drawn from both semantic
categories with equal probability.

Results and Discussion

In this experiment we found that semantic context could be used to mitigate PI if the semantic context of the past trial was different than the current trial. Additionally, we also obtained a reliable SI effect in the recent probes task, which allowed us to compare and contrast both effects within the same experimental paradigm.

Mean accuracy and median response times (RTs) for correct responses to each probe type are presented in Figure 4.4. In order to assess the effects of Semantic Context and Probe Recency on accuracy and RT for negative probes, we conducted a 2x2x2 mixed effects ANOVA including within subjects variables Semantic Context (mixed vs. mismatch), and Recency (recent vs. non-recent), and with Task Order as a between subjects variable (we analyzed the match trials separately, because there were no RN match trials in this experiment). We found a main effect for Recency on both accuracy, $F=5.31$, $p<.05$, $\eta2=.40$, and RT, $F=35.01$, $p<.001$, $\eta2=.36$, indicating that, collapsing over contexts, participants were slower and less accurate in their responses to recent vs. non-recent negative probes (a replication of the standard recency effect). The main effect for Semantic Context was reliable with respect to accuracy, $F=7.45$, $p<.05$, $\eta2=.48$, indicating increased accuracy on mismatch relative to mixed probes (see Fig. 3). There was also a trend toward decreased RTs for mismatch vs. mixed probes, $F=7.45$, $p<.07$, $\eta2=.36$, though this main effect did not reach statistical significance.

Critical to our question of whether semantic information can be leveraged to reduce PI, we found a reliable Semantic Context by Probe Recency interaction with respect to RT, $F=5.68$, $p<.05$, $\eta2=.42$, indicating a decrease in PI in the mismatched context. A similar interaction pattern was shown for accuracy at a marginal level,

F=4.40, p<.07, $\eta2$=.36. In order to directly assess the extent to which the semantic

context of RN probes changed the amount of time required to correctly reject them, we

conducted planned comparisons between probe accuracy and RT for RN_mismatch

versus RN_mixed probes. Results showed increased accuracy, t=2.77, p<.05, *d*=.97, and

reduced RT, t=-2.65, p<.05, *d*=.64, for RN_mismatch probes relative to RN_mixed

probes (see Figure 3). These results indicate a reduction in interference when RN probes

were drawn from a semantic category that differed from that present in the memory set of

the current trial, and suggest that, as predicted, semantic information can be used to

mitigate the effects of PI. Interestingly, although interference was decreased, PI was still

evident in RTs for these trials, with participants taking an average of 41ms (SE=11.34)

longer to correctly reject a RN_mismatch vs. NRN_mismatch trials (t=3.62,p<.05,

d=.50). This indicates that even when the semantic context of an RN probe is

incongruent with that of the current memory set, previously studied information

continues to induce interference.

In order to assess the effects of semantic familiarity independently from PI, we

compared RT and accuracy for NRN_match and NRN_mismatch trials using paired t-

tests. Both SI measures were reliable, t=5.31, p<.001, *d*=.57 for RT and t=2.89, p<.05,

d=.92,for accuracy. This indicates that semantic interference was reliably induced by

NRN probes that were drawn from the same semantic context as memory set items.

In summary, Exp. 2 showed significant reductions in PI when RN probes were

drawn from a different semantic category than memory set items. This suggests that the

semantic context of a recent negative probe can be used to mitigate PI. On both mixed

and mismatch trials, recent negative probes were those that have been studied recently,

and therefore should be activated more than non-recent negative probes. However, on mismatch trials, the semantic context of the current memory set was incongruent with the recent negative probe. One possibility, consistent with our interpretation of effects in Exp.1, is that the semantic tags are used as an additional source of information that may boost the signal to noise ratio between NRN and RN probes on mismatch trials (Nairne, 2002). When the semantic context of each trial is predictable, the status of a recent negative probe is more efficiently resolved when this probe does not match the semantic context relevant for the current trial. Intriguingly, although the magnitude of PI was reduced on mismatch trials, it was not entirely eliminated. This suggests that recent exposure to an item produces reliable interference even when this item could be rejected based solely on its distinct semantic context.

EXPERIMENT 3

Exp. 2 demonstrated that PI is reliably reduced when the semantic context of a RN probe is incongruent with that of the current memory set. One explanation of this finding is that contextual cues contribute to the process by which the familiarity of a recent probe is adjudicated (Nairne, 2002). Thus, when semantic cues associated with current memoranda are sufficiently distinct from those associated with RN probes, interference is reduced.

This interpretation suggests that if the semantic context between current memory set and a RN probe overlaps, interference should increase due to the need to negotiate both the semantic and temporal familiarity of this item. Exp. 3 was designed to examine this possibility. We did this by incorporating RN probes from the same semantic

category as memoranda into Categorized RP task (Exp. 2). This allowed for the assessment of two forms of RN trials: RN_match trials, in which the RN probe matched the semantic category of the current memory, and RN_mismatch trials, in which the RN probes where from a different semantic category. This design allowed us to examine interference caused by both temporal and semantic familiarity using an additive factors framework (Sternberg, 1969). We predicted that if semantic tags are utilized in the assessment of recent negative probes, we should find a recency by semantic context interaction, such that interference is increased for RN_match probes and reduced for RN_mismatch probes.

Method

Participants

Twenty-four participants (10 males; mean age 22.5 years) were recruited from the University of Michigan to participate in the study. All participants gave informed consent as reviewed by the University of Michigan's Institutional Review Board. Participants were paid $15 per hour for their participation plus bonuses for fast and accurate responding throughout the experiment. Bonus scores were calculated as in Exps. 1 and 2.

Procedure

Exp. 3 employed the same Mixed RP task as Exp. 2 (Figure 4.3A). We made a single change to the Categorized RP task: rather than alternating predictably, memory sets in the CAT version of the task varied randomly (see Figure 4.5). This change made it possible for the same semantic category to appear on multiple trials in a row. As such,

recent negative probes could be members of the same semantic category of the current
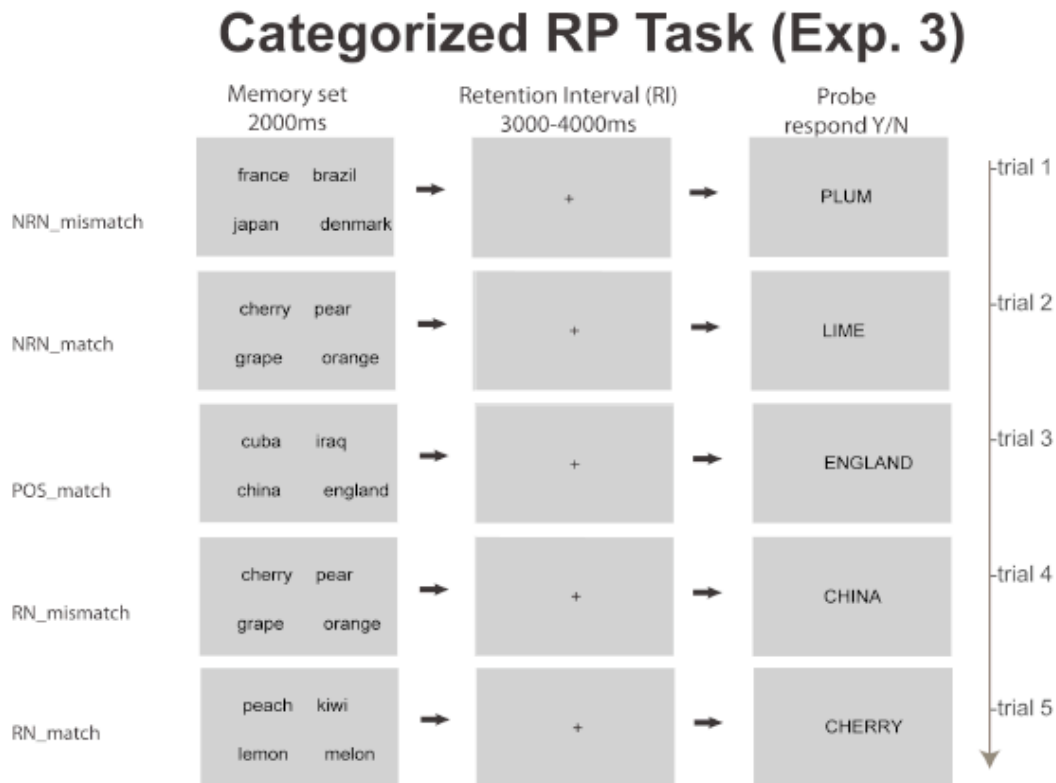
memory set, or a semantically different category.

Categorized RP Task (Exp. 3)

	Memory set 2000ms	Retention Interval (RI) 3000-4000ms	Probe respond Y/N	
NRN_mismatch	france brazil japan denmark	+	PLUM	⊢trial 1
NRN_match	cherry pear grape orange	+	LIME	⊢trial 2
POS_match	cuba iraq china england	+	ENGLAND	⊢trial 3
RN_mismatch	cherry pear grape orange	+	CHINA	⊢trial 4
RN_match	peach kiwi lemon melon	+	CHERRY	⊢trial 5

Figure 4.5.
The Categorized RP task utilized Exp. 3. Rather than alternating from trial to trial (fruit-country-fruit-country), as in Exp. 2, the category membership of memory set stimuli in Exp. 3 was unpredictable, such that adjacent trials could be drawn from either the same category (fruit-fruit) or different categories (fruit-country). This change allowed us to examine 2 types of RN and NRN probes. RN_match probes were recently studied items from the same category presented in the current memory set; RN_mismatch probes were recently studied items from a different category than that presented in the current memory set. Example trials are shown above.

To accommodate these new RN trials, we increased the number of total trials to

256. In the MIX task, there were 128 POS_mixed trials, 64 RN_mixed trials and 64

NRN_mixed trials. In the CAT task there were 128 POS_match trials, 32 NRN_match

trials, 32 NRN_mismatch trials, 32 RN_match trials and 32 RN_mismatch trials. Across

all trials in the CAT task, half of all memory set items and half of all probe items were

95

drawn from each semantic category. As in Exp. 2, the same words were used in the MIX and CAT tasks; memory sets in the MIX task contained a mixture of fruit and country words, and probes were drawn equally from both semantic categories.

Results and Discussion

In Exp. 3 we found interactive effects between SI and PI. When RN probes were taken from the same semantic category as the current stimulus set we found that SI and PI interacted reliably. In addition, we replicated our findings from Exp. 2, where semantic context could reliably mitigate PI if RN probes were drawn from a different semantic category than the current memory set.

Mean accuracy and median RTs for correct responses to each probe type in each task are presented in Figure 4.6. Repeated measures ANOVAs were conducted to assess the effects of Recency (recent vs. non-recent), Semantic Context (mixed, mismatch or mixed), and Task Order (MIX first vs. CAT first) on probe accuracy and RT. With respect to negative probes, we found main effects of both Recency and Context on probe accuracy (F=32.94, $\eta2$=.60, and F=16.86, $\eta2$=.43, respectively, p<.001 for both) and on RT for correctly rejected negative probes (F= 51.03, $\eta2$=..70, and F=10.39, $\eta2$=.32, respectively, p<.001 for both). There was no main effect of Context on positive probe accuracy or RT, indicating that participants were equally fast and accurate at recognizing memoranda from both categorized and mixed stimulus sets.
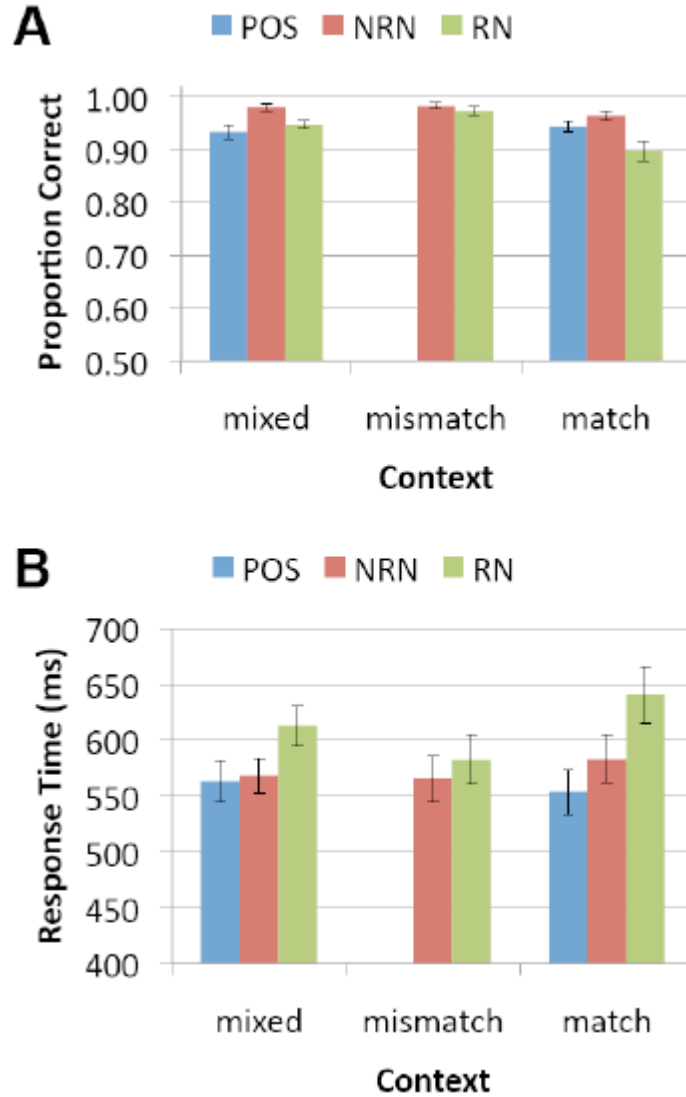
96

Figure 4.6.
Mean accuracy and median RTs from Exp. 3. On mixed trials, which were
blocked, memory sets included a mixture of words from two semantic categories
(fruits and countries), and recent negative probes were drawn from each category
with equal probability. On both mismatch and match trials (which appeared in the
categorized block), memory sets contained items from only one semantic category. In
Exp. 3, the semantic category of memory sets in the categorized block alternated
randomly. Both recent negative (RN) and non-recent negative (NRN) probes were
drawn with equal probability from either the same category as current memoranda
(match trials), or from the other category (mismatch trials).

Critically, we found a Context by Recency interaction on both independent measures, F=9.92, $\eta2$=.31, for accuracy and F=7.86, $\eta2$=.26, for RT, p<.001 for both, indicating that PI reliably varied between Mixed, Matched, and Mismatched semantic contexts. In order to test our hypothesis that interference would increase when RN probes were drawn from the same semantic category of the current memory set, we conducted planned comparisons for RT and accuracy on RN_match vs. RN_mixed trials. Findings showed a significant 5% reduction in probe accuracy, t=-3.96, p<.001, *d=.74,* and a 27.5ms increase in RT, t=2.23, p<.05, *d*=1.29, associated RN_match vs. RN_mixed probes. We also replicated findings from Exp. 2, demonstrating decreased interference when RN probes were drawn from a different semantic category than memoranda. Participants were 3% more accurate, t=2.63, p<.05, *d*=.63, and 30.5ms faster, t=2.90, p<.05, *d*=.31, for RN_mismatch vs. RN_mixed probes.

As in Exp. 2 we assessed the influence of SI independently from PI by comparing RT and accuracy for NRN_match and NRN_mismatch trials using paired t-tests. Results replicated those findings, showing that NRN probes from the same semantic category were more error-prone, t=-2.28, p<.05, *d*=.53, and took longer to reject, t=2.27, p<.05, *d*=.18, than NRN probes from a different semantic category.

Taken together, results from Exp. 3 suggest that the semantic relationship between RN probes and memory set items changes the extent to which these items induced proactive interference. Consistent with predictions, we observed an interaction between Probe Recency and Semantic context, such that interference increased when RN probes shared a semantic context with memoranda, and decreased when semantic contexts were distinct. According the to the additive factors approach (Sternberg, 1969),

two manipulations that effect the same stages of cognitive processing should have interactive effects, whereas manipulations that effect separate stages of processing should have additive effects. Within this framework, the observed interactions are consistent with the notion that a shared interference resolution process is recruited to resolve interference due to both temporal and semantic familiarity. RN_match probes, which are both temporally and semantically familiar, induce more interference than RN_mixed probes, which have episodic familiarity but share no informative semantic relationship with the memory set. RN_mismatch probes, which are temporally familiar but semantically distinct from memoranda, induce the least amount of PI.

GENERAL DISCUSSION

To date the effects of proactive and semantic interference in short-term memory have been investigated largely independently from one another. The purpose of the present work was to jointly examine these effects, and determine the extent to which a common process may be recruited to resolve interference induced by both episodic and semantic familiarity. Exp. 1 demonstrated that the requirement to engage in articulatory suppression during the retention interval of an item recognition task increased interference induced by the need to reject both recently studied items (PI), and items which had never been studied, but which were related in meaning current memoranda (SI). That AS increased both SI and PI in similar ways suggests that some shared cognitive processes may underlie the resolution of both SI and PI. Furthermore, this work replicates our previous findings (Atkins & Reuter-Lorenz, 2008) by showing semantic interference can be induced without a build-up of PI, demonstrating reliable SI

99

using a paradigm in which lure probes were never studied and in which the semantic theme of each memory set was unique.

Exps. 2 and 3 specifically investigated interactions between PI and SI by incorporating both episodic and semantic familiarity into a single paradigm. Consistent with our predictions, we found an interaction between the semantic context of a recent negative probe and the amount of interference produced. This suggests that our vulnerability to PI changes depending on the meaningful relationship between previously studied information and items currently in memory. Interestingly, PI was reliably reduced when recent negative probes were drawn from a semantic context distinct from memoranda, suggesting that semantic cues can be leveraged to reduce interference from recently studied items. Despite this reduction however, significant residual PI remained, suggesting that even when recent negative probes could feasibly be rejected based on semantic information alone, this was not sufficient to eradicate the pervasive effects of episodic familiarity.

The reduction in PI induced by RN probes from a semantic context distinct from that of the memory set was complimented by an increased in PI when RN probes were drawn from the same semantic category. This interaction of semantic context and PI suggests that a common cognitive mechanism is recruited to resolve interference induced by both temporal and semantic familiarity, and offers support for cue-based and feature-based unitary models of memory such as that proposed by Nairne (2002). This model accounts for the pattern of interference effects we observe by predicting an increase in interference as cues become less distinct. For probes related in meaning to memoranda, there is overlap in the semantic cues associated with each, whereas for RN probes, there

is overlap in episodic cues. When an RN probe shares a semantic context with memoranda, both semantic and episodic cues overlap, leading to increased interference. By the same token, when contextual cues become more distinct, interference is mitigated.

The notion that semantic and episodic familiarity are resolved by common mechanisms is also consistent with a recent neuroimaging study, which showed overlap between neural mechanisms of semantic interference in a release from PI paradigm and those previously shown to mediate familiarity-based PI (Oztekin, et al., 2008). Furthermore, ongoing work in our lab (Atkins & Reuter-Lorenz, in prep.) suggests common frontal regions are recruited in response to interference induced by the RP and ST-DRM tasks used here.

The present work offers a unique contribution to the literature by showing that semantic context can change the degree to which we are vulnerable to PI, increasing this vulnerability when semantic context overlaps, and decreasing (though not eliminating) it when semantic contexts are distinct. Taken together, these findings are consistent with the notion that a shared cognitive control is recruited in service of resolution of multiple forms of interference (Oztekin, et al., 2008; Nelson, et al., 2009). Furthermore, this work may contribute to the burgeoning field of research focused on the development of cognitive training regimens that may reduce interference-induced memory distortions (Persson & Reuter-Lorenz, 2008). Identification of the circumstances in which interference is mitigated and enhanced may be crucial to both the development of improved interventions, as well as our increased understanding of the cognitive control of memory.

ACKNOWLEDGEMENTS

REFERENCES

Atkins, A. S., & Reuter-Lorenz, P. A. (2008). False working memories? Semantic distortion in a mere 4 seconds. *Mem Cognit, 36*(1), 74-81.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589.

Badre, D., & Wagner, A. D. (2005). Frontal Lobe Mechanisms that Resolve Proactive Interference. *Cereb. Cortex, 15*(12), 2003-2012.

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In Search of Decay in Verbal Short-Term Memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 35*(2), 317-333.

Collins, A. M., & Loftus, E. F. (1975). A Spreading-Activation Theory of Semantic Processing. *Psychological Review, 82*, 407-428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 240-247.

Conrad, R. (1964). Acoustic Confusions in Immediate Memory. *British Journal of Psychology, 55*(1), 75-&.

Hancock, T., & Hicks, J. (2002). Backward associative strength determines source attributions given to false memories. *Psychonomic Bulletin & Review, 9*(4), 807-815.

Huttenlocher, J., & Kubicek, L. F. (1983). The source of relatedness effects on naming latency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(3), 486-496.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139*(1), 181-193.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences of the United States of America, 95*(14), 8410-8413.

Kincaid, J. P., & Wickens, D. D. (1970). Temporal gradient of release from proactive interference. *Journal of Experimental Psychology, 86*(2), 313-&.

Laughery, K. R., & Harris, G. J. (1970). Visual and auditory intrusion errors in short-term memory. *J Exp Psychol, 83*(1), 101-106.

Levy, B. A. (1971). Role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior, 10*(2), 123-132.

Lewandowsky, S., Geiger, S. M., & Oberauer, K. (2008). Interference-based forgetting in verbal short-term memory. *Journal of Memory and Language, 59*(2), 200-222.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 53*(3), 626-646.

Monsell, S. (1978). Recency, Immediate Recognition Memory, and Reaction-Time. *Cognitive psychology, 10*(4), 465-501.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53-81.

Nee, D. E., Jonides, J., & Berman, M. G. (2007). Neural mechanisms of proactive interference-resolution. *Neuroimage, 38*(4), 740-751.

Nee, D. E., Wager, T. D., & Jonides, J. (2007). Interference resolution: insights from a meta-analysis of neuroimaging tasks. *Cogn Affect Behav Neurosci, 7*(1), 1-17.

Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: evidence for facilitatory and inhibitory processes. *Memory & Cognition, 4*(5), 648-654.

Nelson, J. K., Reuter-Lorenz, P. A., Persson, J., Sylvester, C. Y., & Jonides, J. (2009). Mapping interference resolution across task domains: a shared control process in left inferior frontal gyrus. *Brain Res, 1256*, 92-100.

Nelson, J. K., Reuter-Lorenz, P. A., Sylvester, C. Y., Jonides, J., & Smith, E. E. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proc Natl Acad Sci U S A, 100*(19), 11171-11175.

Oztekin, I., Curtis, C. E., & McElree, B. (2008). The Medial Temporal Lobe and the Left Inferior Prefrontal Cortex Jointly Support Interference Resolution in Verbal Working Memory. *J Cogn Neurosci, Corrected proof available online*.

Persson, J., & Reuter-Lorenz, P. A. (2008). Gaining Control Training Executive Function and Far Transfer of the Ability to Resolve Interference. *Psychological Science, 19*(9), 881-888.

Roediger, H. L., & McDermott, K. B. (1995). Creating false Memories: remembering words not presented in lIsts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803.

Roediger, H. L., McDermott, K. B., & Robinson, K. J. (1998). The role of associative processes in producing false memories. In M. Conway, S. Gathercole & C. Cornoldi (Eds.), *Theories of Memory II* (pp. 187-245). Hove, Sussex: Psychological Press.

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

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' Method. *Acta Psychologica, 30*, 276-315.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*(26), 14792-14797.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of experimental psychology, 70*, 122-129.

Wickens, D. D. (1970). Encoding categories of words - empirical approach to meaning. *Psychological Review, 77*(1), 1-15.

Wixted, J. T. (2005). A theory about why we forget what we once knew. *Current Directions in Psychological Science, 14*(1), 6-9.

Chapter V

CONCLUDING REMARKS

Summary of findings

The present dissertation examined distortions of memory that occur within seconds of encoding. In Chapter II, I provide evidence for false recall and recognition of unstudied items related in meaning to recently encoded memoranda. These findings highlight the vulnerability of verbatim memory processes, even when rehearsal of memoranda is uninterrupted and retention intervals are short (Exps. 2A and B). The incidence of false recall under these conditions is particularly striking, and suggests that unstudied lure items may sometimes become explicitly activated either at the time of encoding itself, or immediately thereafter (Underwood, 1965; Roediger, McDermott, & Robinson, 1998).

In recognition, correct rejection of lure items was remarkably slower than correct rejection of unrelated negative probes, suggesting that even when lures are not recognized as studied items, they induce interference that must be resolved in the service of accurate task performance. This interference resolution process is further examined in Chapters III and IV, in which I investigate the neural mechanisms associated with this process, and explore the relationship between this form of interference and proactive interference (PI) induced by episodic familiarity.

Neuroimaging results (Chapter III) show that semantic interference (SI) is associated with increased activation in a left mid-ventrolateral prefrontal cortex (L VLPFC) and a region of right posterior parietal cortex (R PPC) along the superior boarder of the intra-parietal sulcus (IPS). Region of interest (ROI) analyses of brain-behavior relationships in these two regions revealed a positive correlation between individual differences in L VLPFC activation and response time (RT) indices of SI. This finding replicates reports of similar correlations between activity in this region and PI induced by the need to reject a recently studied item that is no longer task relevant (Jonides & Nee, 2006). Comparisons between probe-related activity to lures that were correctly rejected and falsely recognized indicate that L VLPFC responds to the presence of interference in both cases, and does not distinguish between instances in which interference resolution is successful (correct rejections) and when it fails (false alarms).

By contrast, R PPC activity showed a marginal negative correlation between changes in activation for lure vs. negative probes and reductions in SI. This finding suggests this region may play a role in adjudicating the familiarity of lure items in support of timely and accurate task performance. This interpretation gained support from an exploratory functional connectivity analysis demonstrating positive correlations between R PPC and activations in a network of regions associated with source memory and decision making under uncertainty. Taken together, these findings suggest a possible distinction between regions that respond to the detection of interference in our environment and those that support accurate task performance the face of this interference.

The observation that L VLPFC responds similarly to SI and PI is consistent with the hypothesis that a common cognitive process may be recruited for interference resolution associated with both semantic (SI) and episodic (PI) familiarity. Chapter IV describes a series of investigations that test this hypothesis by examining the relationship between these two forms of interference in short-term memory. Results show an increase in both SI and PI resulting from the need to engage in articulatory suppression during the retention interval. Furthermore, the semantic context of recently studied items predicts the extent to which these stimuli induce interference. When recent negative items are drawn from a semantic category that is distinct from that associated with the current memory set, PI is reduced. This suggests that semantic information can be leveraged to reduce interference in these cases. When recent negative probes are drawn from the same semantic category as current memoranda, PI is increased, suggesting over-additive effects of semantic and episodic familiarity.

Further considerations

Unitary vs. multi-storage models of memory

Although the current investigations cannot directly distinguish between unitary and multi-storage models of memory (see Chapter I), they are consistent with multi-storage models that include components that allow for interactions between representations in short and long-term memory. For example, in a revised version of his working memory model, Baddeley (2000) proposed the episodic buffer, a slave to the central executive, that contains multi-dimensional representations of knowledge gathered from several sources. These sources include representations from episodic memory, as

well as the contents of the phonological and visuospatial buffers. The episodic buffer accounts for influences of episodic and semantic memory on short-term remembering by providing a point of cross-talk between architecturally separate short- and long-term memory systems.

Although this is a reasonable account of semantic influences on short-term memory, evidence in support of the episodic buffer has been elusive (see Jonides, Lewis, Nee, Lustig, et al., 2008). By proposing a replication of long-term memory representations in the episodic buffer, the account lacks parsimony. Furthermore, Baddeley (2000) assumes that conscious awareness is associated with retrieval from the buffer. This claim is difficult to reconcile with findings that PI is evident even when subjects are not consciously aware of manipulations of probe recency, and continues to occur when subjects are explicitly instructed to ignore information from past memoranda (Berman, Jonides, & Lewis, 2009). These findings suggest that retrieval of long-term memory representations need not be accompanied by conscious awareness, and that such representations can implicitly interfere with verbatim retrieval from short-term memory.

Given these considerations, results from the present investigations may be more parsimoniously accounted for by unitary feature-based accounts of memory such as Nairne's (2002). According to this model, what is retained over the short- and long-term is not a particular series of items, but a set of activated features corresponding to task-relevant dimensions of memoranda. Overlap between features associated with memoranda and those activated at retrieval results in interference, and increases the probability of memory errors. An overlap between semantic features thus results in SI,

while an overlap in episodic features results in PI. When both semantic and episodic features overlap, interference is maximized (Chapter IV).

Another reason to favor this unitary-store account pertains to the localization of memory processes in the brain. Baddeley's revised multi-store model (Baddeley, 2000) remains inconsistent with findings demonstrating considerable overlap between the neural mechanisms supporting short- and long-term remembering (Cabeza, Dolcos, Graham, & Nyberg, 2002; Ranganath & Blumenfeld, 2005; Karlsgodt, Shirinyan, van Ep, Cohen, & Cannon, 2005; Ranganath & Blumenfeld, 2005). By assuming a single set of memory representations, the unitary activation-based account is consistent with the notion that a common neural system is recruited to access these representations in both short- and long-term memory.

Phenomenology

Throughout the present dissertation, I use the term *false memory*, to describe erroneous recognition and recall of unstudied items that share a meaningful relationship with recently encoded memoranda. I often employ this term in favor of *false alarm* and *intrusion* in order to highlight the similarity of these effects to semantic distortions that have been documented in long-term memory (see Chapter I). There is an outstanding question, however, regarding the phenomenological experience associated with false recall and recognition that occurs only seconds post-study. Investigations using standard Deese-Roediger-McDermott (DRM) paradigms (Deese, 1959; Roediger & McDermott, 1995) have indicated that false recall and recognition are often associated with strong

subjective feelings of memory, and that these errors are often made with high levels of confidence (Gallo, 2006).

Recent investigations in our lab suggest that similar phenomenological feelings of remembering accompany false recognition in our ST-DRM paradigm (Flegal, Atkins, & Reuter-Lorenz, in preparation). For instance, Flegal et al. (in prep.) examined confidence and *remember/know* judgments associated with false recognition that occurred either 4 seconds or 20 minutes following encoding 4-item DRM memory sets. Results showed that subjects were equally confident in their false recognitions of lures across both delays. Furthermore, rates of *remember* judgments, which connote strong feelings of subjective recollection, were equivalent for lure probes presented following long and short retention intervals. These findings suggest that the short-term memory distortions described in the present dissertation are associated with the similar subjective experiences of remembering, and bolster support for the notion that common associative processes of memory contribute to semantic distortions that occur over short and long retention intervals.

Closing remarks

The present work demonstrates the surprising limitations of verbatim memory processes, and shows that memory is vulnerable to considerable distortion within a very short timeframe. This work was equally motivated by investigations demonstrating false memory in the episodic domain (Loftus & Pickrell, 1995; Roediger & McDermott, 1995; Roediger, 1996; Loftus, 1997, 2003; Schacter, Chiao, & Mitchell, 2003; Schacter & Slotnick, 2004; Gallo, 2006), and findings concerning interference resolution and

cognitive control in short-term retrieval (Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Nelson, Reuter-Lorenz, Sylvester, Jonides, & Smith, 2003; Jonides & Nee, 2006; Badre & Wagner, 2007). Taken together, these findings suggest that short-term remembering is vulnerable to semantic distortion that mirrors false memory effects described in relation to episodic and autobiographical memory. Furthermore, there may be considerable overlap between the cognitive and neural mechanisms that allow for the successful resolution of interference induced by semantic and episodic similarity across a variety of tasks. Future work will serve to elucidate our understanding of the circumstances in which we are most vulnerable to interference and memory distortion. Such investigations may allow for the development of more accurate models of memory and performance in the real world, and may contribute to the design of cognitive interventions to reduce the extent to which this vulnerability leads to errors in judgment and action.

REFERENCES

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901.

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In Search of Decay in Verbal Short-Term Memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 35*(2), 317-333.

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *Neuroimage, 16*(2), 317-330.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17-22.

Flegal, K. E., Atkins, A. S., & Reuter-Lorenz, P. A. (in preparation). False memories seconds later: The rapid and compelling onset of illusory recognition.

Gallo, D. A. (2006). *Associative Illusions of Memory: False Memory Research in DRM and Related Tasks*. New York: Psychology Press.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annu Rev Psychol, 59*, 193-224.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience, 139*(1), 181-193.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences of the United States of America, 95*(14), 8410-8413.

Karlsgodt, K. H., Shirinyan, D., van Ep, T. G. M., Cohen, M. S., & Cannon, T. D. (2005). Hippocampal activations during encoding and retrieval in a verbal working memory paradigm. *Neuroimage, 25, 25*(4), 1224-1231.

Loftus, E. F. (1997). Creating false memories. *Scientific American, 277*(3), 70.

Loftus, E. F. (2003). Make-believe memories. *The American Psychologist, 58*(11), 867.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25*(12), 720-725.

Nelson, J. K., Reuter-Lorenz, P. A., Sylvester, C. Y., Jonides, J., & Smith, E. E. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proc Natl Acad Sci U S A, 100*(19), 11171-11175.

Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences, 9*(8), 374-380.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language, 35*(2), 76-100.

Roediger, H. L., & McDermott, K. B. (1995). Creating false Memories: remembering words not presented in lIsts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803.

Roediger, H. L., McDermott, K. B., & Robinson, K. J. (1998). The role of associative processes in producing false memories. In M. Conway, S. Gathercole & C. Cornoldi (Eds.), *Theories of Memory II* (pp. 187-245). Hove, Sussex: Psychological Press. .

Schacter, D. L., Chiao, J. Y., & Mitchell, J. P. (2003). The seven sins of memory: implications for self. *Annals of the New York Academy of Sciences, 1001*, 226.

Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron, 44*(1), 149.

Underwood, B. J. (1965). False recognition produced by implicit verbal responses. *Journal of experimental psychology, 70*, 122-129.