2011

Rethinking Buffer Operations in a Dual-Store Framework

Melissa Lehman
University of South Florida, mlehman@mail.usf.edu

Rethinking Buffer Operations in a Dual-Store Framework

by

Melissa Lehman

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Psychology
College of Arts and Sciences
University of South Florida

Major Professor: Kenneth Malmberg, Ph.D.
Jonathan Rottenberg, Ph.D.
Mark Goldman, Ph.D.
Cathy McEvoy, Ph.D.
Joseph Vandello, Ph.D.

Date of Approval:
March 29, 2011

Keywords: episodic memory, context, directed forgetting, memory models, serial position effects

Acknowledgements

First and foremost, I would like to thank my advisor, Ken Malmberg, for his guidance and expertise, for teaching me how to do good research, and for drilling into my awareness the idea that data always comes first! I have learned more than I could possibly explain in my time as a graduate student in your lab. Thanks also for your flexibility, for giving me the freedom to develop new lines of research while continuing to guide me toward my goal. I am extremely grateful for the opportunity to have such an experience as a graduate student.

Thanks to Jon Rottenberg, Cathy McEvoy, Mark Goldman, Joe Vandello, and Doug Nelson for taking the time to serve on my committee, and for putting much thought into improving my work. Your guidance has been invaluable to me in this process. I would also like to thank Dewey Rundus for chairing my defense. Thanks also to Judy Bryant for all of the work you have put into making our department a pleasant and productive environment for graduate students, and for your general guidance about research, program requirements, and the academic career. Thanks to all of my friends and colleagues in the psychology department, who have provided me with thoughtful conversation and many encouraging words.

Thanks to my family, none of whom understand why I would want to go to school for so long, but who supported me with pride the entire time anyway. Finally, I want to thank my husband, Erik Lundberg, for being caring, helpful, and supportive of all of my ventures, both in and outside of academia. Thanks for helping me with mathematical formulas whenever I need to use them, for keeping me sane when I am overwhelmed, for tolerating my quirks, and for always putting the toilet seat down.

Table of Contents

List of Tables

List of Figures

Abstract

Atkinson and Shiffrin's (1968) dual-store model of memory includes a structural memory store along with control processes conceptualized as a rehearsal buffer. I present a variant of Atkinson and Shiffrin's buffer model within a global memory framework that accounts for findings previously thought to be difficult for it to explain. This model assumes a limited capacity buffer where information is stored about items, along with information about associations between items and between items and the context in which they are studied. The strength of association between items and context is limited by the number of items simultaneously occupying the buffer. New findings that directly test the buffer assumptions are presented, including serial position effects, and conditional and first recall probabilities in immediate and delayed free recall, in a continuous distractor paradigm, and in experiments using list length manipulations of single item and paired item study lists. Overall, the model's predictions are supported by the data from these experiments, suggesting that control processes, conceptualized as a rehearsal buffer, are a necessary component of memory models.

Chapter 1

Introduction

Any introductory cognitive psychology textbook will likely include a distinction between short-term memory and long-term memory. Short-term memory is often described as a mechanism that can hold a limited amount of information for a short time, from which that information will escape if strategies are not used to maintain it, and long-term memory as a more permanent memory store which can hold vast amounts of information (Goldstein, 2008). Theories that distinguish two types of memory stores date back to William James. In *Principles of Psychology* (1890), James describes *primary memory*, the temporary knowledge of an initial encounter with a stimulus which must be attended to in order to maintain it in consciousness, and *secondary memory*, the knowledge of a stimulus one has experienced before which has since been dropped from consciousness. James further specified that primary memory is an accurate representation of events that occurred, whereas secondary memory is subject to distortion. During what is often called the “Cognitive Revolution” in psychology, Broadbent (1958) devised an information processing model of attention and memory, which includes an immediate memory system and a more permanent long-term module where learned information is stored. Thus, dual-store theories of memory, those positing separate and distinct short-term and long-term memory stores, enjoy a long history in psychology research.

Waugh and Norman (1965) expanded on the primary/secondary memory distinction in an initial attempt to develop a formal model of these processes. The Atkinson and Shiffin (1968) dual-store model of memory, building on this same framework, is perhaps the most commonly cited formal model of memory, often referred to as "the modal model." According to the Atkinson and Shiffin (A-S) model, memory consists of not only structural components, including short-term and long-term stores, but also of control processes that allow an individual to focus attentional resources on to-be-remembered stimuli in order to increase the likelihood of encoding those stimuli (or, increase the amount of information that is encoded about those stimuli). These processes are under the control of an individual, and they influence the way that memories are stored. Of particular interest, their model includes a *rehearsal buffer*, a flexible control process which an individual may use in order to encode information relevant to a given task. The buffer was designed to account for the rehearsal of items during the study of a list of items. According to the A-S model, the rehearsal buffer is a limited capacity system that allows for the temporary storage of information; it is limited in terms of the number of items it can simultaneously accommodate, and when the capacity is reached, an item must be dropped. The model assumed that the number of trials for which an item resides in the rehearsal buffer is positively correlated with its encoding in long-term memory.

Though dual-store models have received much empirical support (Atkinson & Shiffrin, 1971), controversy surrounds the nature of these memory processes. Theorists in the dual-store camp point to evidence from examinations of serial position effects, in addition to findings of differential deficits in short-term and long-term memory in neuropsychological case studies, in support of dual store models (Atkinson & Shiffin, 1968; Cowan, 1995; Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Glanzer & Cunitz, 1966; Waugh & Norman, 1965). On the other hand, others have criticized dual-store models and developed single-store models that are able to account for findings said to be troublesome for dual-store models (Bjork & Whitten, 1974; Craik & Watkins, 1973; Crowder, 1989; 1993; Howard & Kahana, 1999; Nairne, 1996).

The purpose of the current project is to explore the importance of buffer processes typically associated with dual-store models. It should be noted that the buffer operates independently of any structural components of the model and should not be confused with the short-term store. Atkinson and Shiffrin (1968) explicitly stated that the buffer is a process under the control of a subject, used in an attempt to maximize performance. Thus, we focus not on the issue of whether short-term memory is a structural component of the memory system, but whether control processes that allow an individual to manipulate to-be remembered information in a way that is optimal for a given task are necessary for models to account for a variety of memory phenomena.

In addition to exploring the function of the buffer as a memory process, we also wish to revisit another area from classic memory research by exploring the process of *chunking*, combining small units of information into larger meaningful units of information in order to increase the number of items that can simultaneously be maintained by the memory system (Miller, 1956), into the framework of this model. Ideas related to chunking also date back to James (1890), who suggested that one can attend to an indefinite number of *things* and that the number of things that can be attended to depends on the nature of those things. Miller proposed that information can be better remembered if it is of a type that can be chunked together. Theories that grouping information together into meaningful units will allow for remembering more information have a long history of empirical support, and may be consistent with intuitive notions about memory (Gobet, Lane, Croker, Cheng, Jones, Oliver, and Pine, 2001); thus there is little controversy surrounding the issue of whether chunking serves as a memory process. However, chunking does not play a large role in most formal memory models. Aside from Murdock's (1995) description of the role of chunking in serial order effects, chunking processes are rarely described in formal models of episodic memory (others have developed computational models of chunking in learning new information, though these are quite different from models of episodic memory; see Gobet et al., 2001, for a description of such models). Thus, our goal is to explore the interaction of chunking and buffer operations in the framework of a formal memory model.

In sum, the broad intent of this project is to outline a formal computational model of memory that accounts for a variety of patterns of data in different experimental paradigms with very few free parameters. This model will include assumptions based in the A-S buffer model framework and integrate ideas about chunking in the encoding process. The model will generate testable predictions, which can inform us about the nature of human memory in a clear cut manner (Widaman, 2008). In this manuscript, I will address various criticisms of dual store models, beginning with a discussion of such criticisms, followed by a description of the Lehman-Malmberg model, a model initially developed in order to account for directed forgetting effects (Lehman & Malmberg, 2009; 2011; Malmberg, Lehman, & Sahakyan, 2006). I will next discuss the ways in which this model can account for data that have troubled buffer models in the past. Finally, I will present the results of some experiments conducted in order to test the model's predictions in a novel experimental paradigm. Specifically, I will examine whether retrieval from a buffer is a necessary component of the model, or whether the model can account for the data without such a process. Ultimately, I wish to show that a model based in the A-S framework which includes a buffer process is both able to handle data that has been claimed to trouble such models in the past and new data that will be hard for a single-store model that does not include such control processes to account for.

1.1 Serial Position Effects

The A-S model is well known for its account of serial position curves, specifically *primacy* and *recency* effects. Primacy refers to the greater probability of recall for items that come from the beginning of a list, and recency to the greater probability of recall for items that come from the end of a list (Deese & Kaufman, 1957). According to dual-store models, primacy occurs because items from early serial positions occupy a spot in the rehearsal buffer for a longer period of time, whereas recency occurs due to the retrieval of items that occupy the rehearsal buffer at time of test, which is assumed to be highly accurate (Atkinson & Shiffrin, 1968).

Using an overt rehearsal procedure, Rundus (1971) showed that items at the beginning of the list receive more rehearsals throughout the list than items that occur later in the list, consistent with the hypothesis that items at the beginning of the list spend more time in the rehearsal buffer. Additional support for dual-store models came from studies showing that certain factors differentially influence primacy and recency. For example, increases in presentation time of studied items increases recall for items toward the beginning of the list without affecting the recency portion of the list. On the other hand, the addition of a distractor task at the end of a study list, which presumably prevents rehearsal of items in the buffer, eliminates recency without affecting the primacy portion of the list (Glanzer & Cunitz, 1966). Murdock (1962) showed that whereas primacy effects are reduced with increasing list length, recency effects were unaffected by list length, suggesting that items in the recency portion of the list exist in a different state than other items on the list, consistent with the idea that they are present in a buffer at time of test.

Further work has shown other dissociations between primacy and recency effects. For example, whereas recency is seen in immediate free recall, a *negative recency* effect, worse recall for the items at the end of the list compared to the middle of the list, is present in final free recall of items from lists that were previously tested under immediate conditions; however, items at the beginning of the list are not differentially recalled in the final test (Craik, 1970). This suggests that the end of list items are present in a privileged state during immediate testing, when they still reside in the buffer, but these items have not been well encoded in long-term memory, so they suffer on a later recall test. Similarly, proactive interference occurs when memory for new items is negatively affected by preceding items. While items throughout a list are increasingly subject to proactive interference, items from the very end of a list are not affected (Craik & Birtwistle, 1971).

Dual store models are able to account for these distinctions between primacy and recency effects with the use of buffer processes during encoding. It is important to note that encoding in long-term memory was not explicitly described by Atkinson and

Shiffrin, aside from the inclusion of an assumption that information about items and about the context in which the item is studied is represented. Later developments in modeling these encoding processes were made in the framework of SAM, most important for the current purposes, the assumption that associations among items that simultaneously share the buffer are encoded (Raaijmakers & Shiffrin, 1981).

1.2 Criticisms of Buffer Models

Despite a myriad of empirical support for the A-S model, dual-store models have been criticized due to some experimental findings which some have suggested are inconsistent with such models. First, such models are based on assumptions that items that spend more time in the rehearsal buffer will be better encoded in long-term memory: “An item's likelihood of being retained increases in direct proportion to its total presentation time within a list.” (Waugh, 1970, p. 587; see also Atkinson & Shiffrin, 1968). Thus, such models predict that items that are studied for longer periods of time should be better remembered, regardless of whether the longer study time is due to increases in presentation time or increases in the number of presentations of the items.

While increases in study time have been shown to be related to greater recall (Glanzer & Cunitz, 1966; Murdock, 1962), this is not always the case. For example, Craik and Watkins (1973) reported that recall did not increase linearly with study time when maintenance rehearsal was used, specifically “neither the length of an item's stay in short-term storage nor the number of overt rehearsals it received was related to subsequent recall” (Craik & Watkins, 1973, abstract).

Additional criticisms of buffer models have been made in response to findings of recency effects in a continuous distractor task, where a short distractor task is presented after each item on a list (Bjork & Whitten, 1974). Despite the distractor task presented after each item, including the last item, the last items on the list are better recalled than earlier items, a finding referred to as *long-term recency* (Bjork & Whitten, 1974; Howard & Kahana, 1999). Critics of buffer models suggest that, as the distractor task should eliminate the most recent items from the buffer, a long-term recency effect in a continuous distractor task is hard for a buffer model to explain (Bjork & Whitten, 1974; Crowder, 1989; 1993; Howard & Kahana, 1999).

Another challenge to buffer models came from findings of contiguity effects in continuous distractor tasks (Howard & Kahana, 1999). Contiguity effects refer to findings that during recall, items from nearby serial positions tend to be output successively. Kahana (1996) examined these conditional response probabilities (CRPs) and found that in a standard free recall paradigm, transitions were most likely to be made to items that were presented in the closest temporal proximity to a recalled item, and that transitions were more likely to be made in the forward direction than the backward direction. These patterns are referred to as *lag-recency effects*. Howard and Kahana (1999) suggested that CRP functions in a continuous distractor task can be useful in distinguishing different models. They proposed that according to a buffer model, a continuous distractor task should disrupt buffer operations, preventing the encoding of associative information between items; thus nearby transitions should not be more likely to occur than distant transitions. They found contiguity effects in a continuous distractor paradigm similar to those seen in standard free recall.

Howard and Kahana (1999) showed that while a two-store model was unable to account for these data, a single store model was able to account for both lag-recency and long-term recency effects. Later, Howard and Kahana (2002) developed the Temporal Context Model (TCM), which assumes that context fluctuates over time, but that this fluctuation is driven by retrieval of prior contextual states. TCM was able to account for long-term recency effects in a single-store model which does not include a rehearsal buffer, because it assumes that the context during test is most similar to the most recently learned item, and this context is used to probe memory. Further, the model accounts for lag-recency effects because it assumes that the context retrieved from a recalled item during test is used to probe memory, thus items in a similar temporal context (i.e. studied in a nearby serial position) will be most likely to be recalled next (see the discussion for a more detailed description of TCM).

More recent developments in buffer models have been important for addressing some of the problems that have troubled such models in the past. Malmberg and Shiffrin (2005) assumed that the encoding of item information and the encoding of context information followed different time courses. Whereas item information is encoded as long as the item occupies the buffer, a fixed amount of context, or "one shot" of context, is stored in perhaps as little as 1-2 seconds. Thus, increases in study time may increase the encoding of inter-item associative information. Malmberg and Shiffrin implemented the one-shot hypothesis in the REM model. These assumptions may be useful in understanding the results of studies using manipulations of study time in incidental learning tasks.

The Lehman-Malmberg model of directed forgetting fleshed out a set of assumptions concerning the encoding of associations in the REM framework. This model included a limited capacity buffer, where information is stored about items, association between items that simultaneously reside in the buffer, and associations between items and the context in which items are studied. Importantly, this model assumed that the strength of association between items and context, and among different items is limited by the number of items occupying the rehearsal buffer. Lehman and Malmberg (2009) experimented with various versions of the model and concluded that buffer operations were necessary in order to account for serial position effects in directed forgetting.

These more recent developments in buffer models may be useful in accounting for the aforementioned phenomena that have troubled the original Atkinson and Shiffrin model. For example, the one-shot model assumption allows buffer models to account for Craik and Watkins' (1973) findings that maintenance rehearsal does not increase recall (though it should be noted that this claim itself is problematic, as recall did increase with lag in their study, albeit not significantly). Lehman and Malmberg's (2009) assumptions regarding buffer operations in the model may allow the model to account for contiguity effects in continuous distractor tasks. Before describing the ways in which the Lehman-Malmberg model can account for these findings, I will first review the development of the Lehman-Malmberg model and describe the data that it has been able to account for using a similar set of parameter values.

1.3 Directed Forgetting

Our interest in buffer models grew out of a goal to develop a model of intentional forgetting. Intentional forgetting is often studied in the lab using the *list method of directed forgetting*. In this task, two lists are studied and subjects are instructed to remember both lists (the "remember" condition), or they are instructed after studying the first list to forget the first list (the "forget" condition). Contrary to the instructions, *both* lists are tested. Typically in free recall, memory is worse in the forget condition compared to the remember condition for words from the to-be-forgotten list and better for words from the to-be-remembered list (Basden, Basden, & Gargano, 1993; Bjork, 1972, 1978; Bjork & Geiselman, 1978; Bjork, LaBerge, & Legrand, 1968; Bjork & Woodward, 1973; Block, 1971; Geiselman & Bagheri, 1985; Geiselman, Bjork, & Fishman 1983; MacLeod, 1998; MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003; Malmberg, Lehman, & Sahakyan, 2006; Sahakyan, 2004; Sahakyan & Delaney, 2003; Sahakyan & Kelley, 2002; Sheard & MacLeod, 2005). These effects are referred to as the *costs* and *benefits* of directed forgetting, respectively. While the costs and benefits of directed forgetting are fairly typical findings in free recall, they are not always observed. Sahakyan and Delaney (2003) have suggested that the costs may be obtained independently from the benefits, and Sahakyan and Goodmon (2010) observed costs, but not benefits, across five directed forgetting experiments. Additionally, costs and benefits have been shown in recognition testing by some researchers (Lehman & Malmberg, 2009), whereas others have reported partial effects (Benjamin, 2006; Loft, Humphreys, & Whitney, 2008; Sahakyan & Delaney, 2005; Sahakyan, Waldum, Benjamin, & Bickett, 2009) and others have reported no effects (Basden, Basden, & Gargano, 1993; Block, 1971; Elmes, Adams, & Roediger, 1970; Geiselman, Bjork, & Fishman, 1983).

While there are many – not necessarily mutually exclusive – hypotheses about specific intentional forgetting findings, until recently, a comprehensive explanation was lacking. The following section will describe various hypotheses for directed forgetting and review work that has been done to develop a comprehensive account of intentional and unintentional forgetting in both recall and recognition.

1.3.1 Differential Rehearsal

Rehearsal plays a well-documented role in many memory models as a mechanism that maintains an item in an accessible state, thereby also increasing the amount of information that is encoded about that item (Atkinson & Shiffrin, 1968; Rundus, 1971). Rehearsal has also been proposed to play an important role in intentional forgetting. According to the *differential-rehearsal hypothesis* (Bjork, LaBerg, & Legrand, 1968; Sheard & MacLeod, 2005) instructions to forget alter the allocation of limited resources during study, and hence the extent to which some items are encoded. Accordingly, subjects stop rehearsing words from the to-be-forgotten list 1 (i.e., L_1) after the forget instruction is given and devote all further rehearsals to the following list 2 (i.e., L_2).[1] In contrast, subjects in the remember condition covertly rehearse items from *L1* while they

[1]The models that we discuss focus on the list method because it is for this procedure that the interactions between recall and recognition have been observed. For the *item method of directed forgetting*, items are presented with a subsequent cue to remember or forget each item. Recognition and free recall for to-be-remembered words is better than for to-be-forgotten words (Roediger & Crowder, 1972; MacLeod, 1975; Woodward & Bjork, 1971). Thus, the differential rehearsal hypothesis assumes that upon the presentation of the remember instruction subjects engage in an elaborative rehearsal process that is not invoked after the instruction to forget (MacLeod, 1975; Woodward, Bjork, and Jongeward, 1973).

study L_2. This reduces the average number of rehearsals allocated to L_2 items and increases the average number of rehearsals allocated to L_1 items. Because the items on L_1 receive more rehearsals after an instruction to remember compared to L_1 items in the forget condition, they are encoded better, and they are more likely to be remembered. This explains the costs of directed forgetting. Because items from L_2 compete with L_1 items for limited rehearsals in the remember condition, they are remembered worse compared to L_2 items in the forget condition, and this produces the benefits of directed forgetting. Indeed, the instruction to forget affects the form of free recall serial position curves (MacLeod, 1998; MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003; Sheard & MacLeod, 2005). For L_2, there is a pronounced primacy effect in the forget condition, and an almost absent primacy effect in the remember condition. Thus, most of the L_2 benefits are associated with enhanced memory for items in the early serial positions. For L_1, the instruction to forget has a smaller effect on the form of the serial position curves, although performance is greater in the remember versus the forget condition, of course.

The differential rehearsal hypothesis is unlikely to provide a complete explanation of list-method directed forgetting for several of reasons. First, the instruction to remember should enhance memory for L_1 items presented at the end of the list on the assumption that they are the L_1 items given extra rehearsals during L_2 in the remember condition. However, this has not been observed at times (Sheard & MacLeod, 2005), and thus, directed forgetting can be observed even when the recency portion of L_1 is unaffected by the instruction to forget. Next, the differential rehearsal hypothesis predicts that directed forgetting should not be observed when rehearsal is discouraged, but it is (Bjork et al., 1968; Block, 1971 Geiselman, et al., 1983; Sahakyan & Delany, 2005). Moreover, every theory of memory predicts that altering the extent of item encoding, via enhanced rehearsal or other means, should improve both free recall and recognition (cf. Malmberg, 2008). The fact that directed forgetting has been observed rarely for recognition memory is problematic for these models. Acknowledging this, Sheard and MacLeod noted that serial position effects might be smaller for recognition, and thus it might be difficult to observe directed forgetting because prior experimental designs were not suitable for observing reliable effects.

1.3.2 Inhibition

The role of inhibition in episodic memory is under active investigation, most notably as it relates to unintentional forgetting in the domain of retrieval-induced forgetting (Anderson & Bjork, 1994; Norman, Newman, & Detre, 2007). However, the possibility that inhibition is used to intentionally forget has also been investigated; some have proposed that inhibition of to-be-forgotten items produces the costs and a concomitant reduction in interference produces the benefits of directed forgetting (Elmes, Adams, and Roediger, 1970; Weiner, 1968; Weiner & Reed, 1969). For instance, subjects might mentally group the to-be-forgotten and to-be-remembered material separately, and then inhibit the to-be-forgotten set during retrieval (Geiselman, Bjork, & Fishman, 1983). Because they are inhibited, these items create less proactive interference, leading to the benefits.

However, the inhibition hypothesis also has difficulty explaining the null effects of intentional forgetting on recognition. To account for them, sometimes the inhibition hypothesis assumes that recognition testing “releases” the to-be-forgotten items from inhibition (Geiselman & Bagheri, 1985; Basden et al., 1993; MacLeod et al., 2003). This suggestion is circular, and usually there is no evidence that the to-be-forgotten items were ever inhibited to start (Basden, Basden, & Wright, 2003; Bjork & Bjork, 1996; Geiselman et al., 1983). Inhibition accounts are further challenged to explain why some recognition experiments exhibit no costs, and yet the benefits remain (Benjamin, 2006; Sahakyan & Delaney, 2005). That is, under what conditions should a release from inhibition be observed, and under what conditions should a release from inhibition not be observed and why? Last, the inhibition hypothesis should explain how subjects place the traces into two separate sets, and inhibit one set and activate the other. In this sense, inhibition accounts describe the data well, but they do not offer much insight into the operations of memory.

1.3.3 Contextual Differentiation

Changes in context play a primary role in forgetting according to many theories (Dennis & Humphreys, 2001; Estes, 1955; Gillund & Shiffrin, 1984; Humphreys, Bain, & Pike, 1989; Howard & Kahana, 2002; Jang & Huber, 2008; Mensink & Raaijmakers, 1989; Murdock, 1997; Murnane, Phelps, & Malmberg, 1999). As the difference between the context features encoded during study and context cues available at test increases, forgetting increases. According to Sahakyan and Kelley's (2002) variant of the *set differentiation* hypothesis of directed forgetting (Bjork et al., 1968; Bjork, 1970), study involves the storage of information representing the studied items (i.e., item information) and the context in which the items occur (i.e., context information). L_1 and L_2 are associated with an overlapping set of contextual elements (e.g., Estes, 1955; Mensink &

Raaijmakers, 1989). The instruction to forget causes an accelerated change in context between lists, and there is less interference between L_1 and L_2. When recalling from L_2, less interference from the L_1 traces produces the benefits of the instruction to forget. The costs are the result of the relative inaccessibility of an effective L_1 context cue due to the relatively rapid change in context that occurred between the list presentations. This is the *contextual differentiation hypothesis.*

The logic behind the contextual-differentiation hypothesis is derived from the literature on context-dependent memory (e.g., Anderson, 1983; Godden & Baddeley, 1975; Goodwin, Powell, Bremmer, Hoine & Stern, 1969; Eich, Weingartner, Stillmin & Gillin, 1975; Macht, Spear, & Levis, 1977, Murnane et al., 1999; Smith, 1979). Sahakyan and Kelley (2002) compared standard directed forgetting conditions to a between-list context-change condition. In the context-change condition, some subjects were given the remember instruction, followed by an instruction to imagine that they were invisible, in order to create a mental context change. Subjects in the remember-plus-context-change condition performed almost identically to subjects in the standard forget condition – showing both costs and benefits of the context change. A strong prediction of the contextual differentiation hypothesis is that the costs and benefits of directed forgetting are dependent on the ability of the subject to mentally reinstate appropriate context cues at test. Indeed, context effects are eliminated or reduced when appropriate context cues are available for both intentional and unintentional forgetting procedures. For instance, Smith (1979) showed that the mental reinstatement of the environmental context eliminates the costs of context dependent memory. In the intentional forgetting literature, Sahakyan and Kelley (2002) used standard remember and forget conditions, but after studying the second list, half of subjects participated in a context reinstatement procedure. Afterward, subjects in the forget and remember-plus-context-change groups showed reduced costs and benefits compared to the groups that did not receive the reinstatement. Presumably, the remaining costs and benefits are due to the use of some contextual elements found at test. In any case, these findings revealed that context reinstatement has similar effects on intentional and unintentional forgetting.

1.4 A Mathematical Model of Directed Forgetting

Malmberg, Lehman, & Sahakyan (2006) set out to develop a formal model of the contextual differentiation hypothesis for directed forgetting, but noted various inconsistencies between existing data sets and the contextual differentiation hypothesis. One problem for the contextual differentiation hypothesis concerned the lack of a recency effect observed in many free recall directed forgetting experiments. All things being equal, any model of free recall that assumes that temporal context plays an important role during retrieval predicts that L_2 should be better remembered than L_1 in the remember condition because L_1 was learned prior to L_2 (Ebbinghaus, 1885). In contrast, sometimes L_1 is actually remembered better than L_2 in the remember condition (Geiselman, Bjork, & Fishman, 1983; Sahakyan & Kelley, 2002; Sahakyan, 2004). This reversed recency effect, better memory at longer retention intervals, suggested to us that the traditional designs used in list-method directed forgetting confound several variables with list order, such as the location of distractor tasks and the presence of proactive interference, and thus give L_1 an advantage over L_2.

Because the list method usually makes use of only two lists, L_1 and L_2, the effect of L_1 versus L_2 is confounded with presence of an interfering prior list. In addition, subjects usually do not perform a distractor task after L_2. Lack of a subsequent distractor task benefits L_1 because the last items on L_1 maybe rehearsed during L_2 (Peterson & Peterson, 1959; Rundus, 1971). Lastly, subjects are typically asked recall both L_1 and L_2at test simultaneously, which makes it somewhat plausible that output interference explains directed forgetting.

1.4.1 Improved Design

To control for these variables, Malmberg et al. (2006) used a three-list design (cf. Sahakyan, 2004), where only the second and third lists were tested. Thus, the first list was referred to as L_0, and for consistency in making comparisons to previous experiments, the second was referred to as L_1, and the third as L_2. Thus, both L_1 and L_2 were preceded by a prior list. Additionally, a distractor task, traditionally used a means for controlling rehearsals (Peterson & Peterson, 1959), was performed after each list. Finally, subjects recalled one list at time in order to control for output interference; those subjects asked to recall L_1 can do so when not also attempting to recall L_2 items (e.g., Sahakyan & Kelley, 2002). Given these changes, Malmberg et al. predicted better memory for L_2 compared to L_1, in addition to the costs and benefits of directed forgetting.

A second prediction concerned intrusion errors. While intrusion rates are usually very low and hard to investigate, the assumptions of this model did generate a few predictions. Because context at test is more similar to the context of L_2 than to the context of L_1, the number of intrusions from L_2 while trying to recall L_1 should be greater than the number of intrusions from L_1 when trying to recall L_2. Additionally, Malmberg et al. predicted that the context differentiation that occurs as the result of the forget instruction will make intrusions less likely in the forget condition.

1.4.2 Mathematical Model

Malmberg et al. (2006) developed a model for the context differentiation that occurs with directed forgetting in a free recall paradigm, based within the framework of the Search of Associative Memory theory (SAM; Raaijmakers & Shiffrin, 1981). According to the SAM model, remembering involves a process of sampling and recovery of stored memory traces (images). Sampling probabilities are determined by the strength of association between contexts and images in memory. This association strength is represented by the parameter, *a*, and is referred to as *context strength*. The parameter, *b*, is the strength of association between two items that were recently rehearsed together when one item is used to cue the image of the other item (*inter-item strength*). The parameter, *c*, is referred to as *self strength*, and it is the associative strength between an item's image and the same item used as a cue. Lastly, the parameter, *d*, is the strength of association between two items that were not recently rehearsed together, and it is referred to as *residual strength*. This is often considered to be a source of noise.

The context strength parameter, *a*, is the key parameter for implementing the context change model in SAM (Shiffrin, Ratcliff, & Clark, 1990; Malmberg & Shiffrin, 2005). For this model, the parameters *b, c*, and *d*, do not influence the model, and thus will be left out of this discussion.

According to SAM, the probability of sampling image, *I*, given *Q* as a retrieval cue is:

$$P(I \mid Q) = \frac{S(I,Q)}{\sum_{J=1}^{m} S(J,Q)},$$

where *m* images are stored, and *S(J,Q)* is the strength of association between the retrieval cue and image, *J = 1…m*. We assumed that each probe of memory is with a context cue only. While SAM assumes that both item and context cues can be used to probe memory, we made the simplifying assumption that *b* is the same for all images, and hence item cues do not differentially affect directed forgetting. Thus, the model attempted to account for directed forgetting using the list method without appeal to a rehearsal account.

Since the list method involves studying more than one list, we assumed that the context changes between them (cf. Mensink & Raaijmakers, 1989). Call the lists L_x and L_y. Given that one is trying to recall items from L_x, there will be an item-to-context association for each L_x image and the context that is used as the retrieval cue (a_{Lx}) and an item-to-context association for each L_y image and context used to probe memory (a_{Ly}). On these assumptions, the probability of sampling image I from L_x is:

$$P(I_{Lx} \mid Q) = \frac{a_{Lx}}{\sum_{J=1}^{m} a_{Lx} + \sum_{J=1}^{m} a_{Ly}}, \qquad (1)$$

and the probability of mistakenly sampling image I from L_y is:

$$P(I_{Ly} \mid Q) = \frac{a_{Ly}}{\sum_{J=1}^{m} a_{Lx} + \sum_{J=1}^{m} a_{Ly}}. \qquad (2)$$

SAM assumes that the process of sampling is noisy and error prone, so we need a recovery process in order to actually retrieve items from memory. Once an image has been sampled from memory a recovery of the contents of that image is attempted, which is successful with the following probability:

$$R(I_n, Q) = 1 - \exp(-S(I_n, Q)) = 1 - \exp(-a_{Ln}) \qquad (3)$$

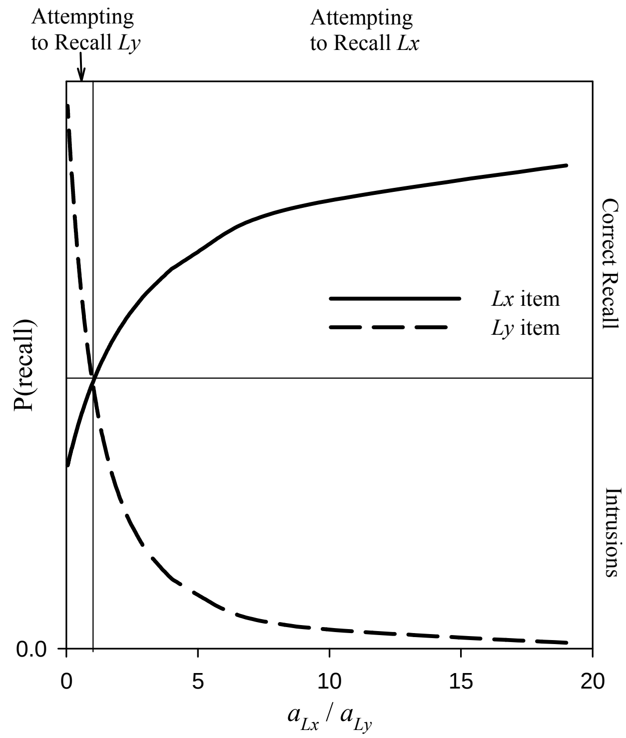
where $n = x$ or y. Thus, the product of Equations 1 and 3 give the probability of successfully recalling a given item from the target list. In contrast, the product of Equations 2 and 3 give the probability of recalling a given item from a non-target list (i.e., an intrusion error).

The SAM model for the context change associated with directed forgetting had three necessary assumptions:

1. $a_{Lx} \neq a_{Ly}$. Context is assumed to change from list to list, and a different context cue is used when attempting to recall L_x versus L_y items. For the sake of simplicity we assumed that context does not change within a given list (cf. Mensink & Raaijmakers, 1989). Thus, the strength of the context-to-image association differs between images on L_x and L_y depending on whether one attempts to recall L_x or L_y. When we want to recall from L_x we use the L_x context cue. When successfully recalling items from L_x, $a_{Lx} > a_{Ly}$, and when successfully recalling items from L_y, $a_{Lx} < a_{Ly}$. This assumption allows the model to predict that it is possible to recall items from a specific list, and that when attempting to recall from a specific list, items from that list are more likely to be sampled than items from another list.

Figure 1 illustrates the relationship between the ratio of a_{Lx} to a_{Ly} and recall performance. When one attempts to recall L_x, the probability of recalling an L_x item

Figure 1. The SAM model of directed forgetting.



Note: a_{Ly} was set to .05 and a_{Lx} was varied from .001 to 1.0

increases and the probability of recalling an L_y item decreases as a_{Lx} / a_{Ly} increases. When one attempts to recall L_y, the probability of recalling an L_y item increases and the probability of recalling an L_x item decreases as a_{Lx} / a_{Ly} decreases.

2. The more recent the list, the greater the strength of the context-to-image association is. The context at test will be more similar to the context of the most recent lists than to the context of earlier lists. Thus, more recent lists should be better recalled than less recent lists. This is implemented by assuming that $a_{Lx} < a_{Ly}$ when attempts are made to recall L_x and L_y, respectively. The assumption also leads to the prediction that intrusions from L_2 will be more likely when recalling from L_1 than intrusions from L_1 when recalling L_2.

3. The forget instruction increases the difference between contexts for different lists by decreasing the context strength for the to-be-forgotten list. Specifically, if L_x was studied before L_y, the instructions to forget L_x will decrease a_{Lx}. We assumed that the instructions to forget L_x, however, have no effect on the strength of the association between the test context used and the L_y images in memory. According to the sampling equation mentioned above, when one is trying to recall from L_x, instructions to forget L_x decrease a_{Lx}, thus causing a decrease in the probability of recalling an L_x item. This creates the costs of directed forgetting. On the other hand, when one is trying to recall from L_y, instructions to forget have decreased a_{Lx}, but not affected a_{Ly}, thus the probability of recalling an L_y item increases – creating the benefits of directed forgetting. Additionally, this context differentiation allows the model to predict decreased intrusions in the forget conditions, where the contexts associated with each list have less overlap than in the remember condition.

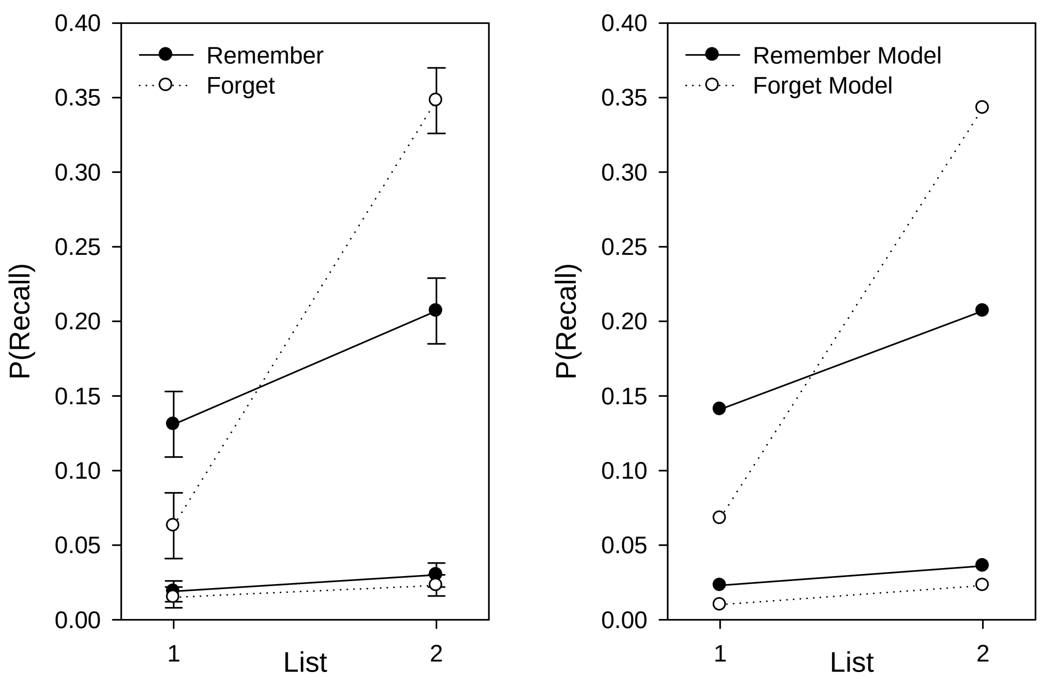
Consider Figure 1. Assume that one is trying to recall L_y. Instructions to forget L_x, that is, decreasing a_{Lx} / a_{Ly}, causes an increase in the probability of recalling an L_y item. This is the benefit of directed forgetting. Now assume that one is trying to recall L_x. Instruction to forget L_x causes a decrease in the probability of recalling an L_x item. This is the cost of directed forgetting.

1.4.3 Experimental Findings

Malmberg et al. (2006) conducted an experiment using the three-list design in order to test the SAM model's predictions in directed forgetting, and found that the model's predictions were supported. The data and model fits are shown in Figure 2.

First, recency was present in both the remember and forget conditions, such that L_2 was better remembered than L_1 in both conditions. Further, the costs and benefits of directed were also present. Finally, while intrusion rates were low and differences were for the most part unreliable, the trends in intrusion rates supported the model's predictions: subjects were more likely to have intrusions from L_2 while being tested on L_1 than they were to have intrusions from L_1 while being tested on L_2, and the probability of either type of intrusion was lower for subjects in the forget condition than in the remember condition.

Figure 2. Data and SAM model predictions for probability of correct recall and intrusion errors for free recall in directed forgetting.



Note. The intrusions in these graphs refer to intrusions that came from either List 1 or List 2. When recalling from List 1, any List 2 item that was output is referred to as an intrusion and vice versa.

Overall, the SAM model provided good fits to the correct recall and intrusion data from a directed forgetting experiment using an improved design to eliminate experimental confounds. The findings support a contextual differentiation model of directed forgetting. Lehman and Malmberg (2009) extended the work of Malmberg et al. (2006) by examining serial position effects and first recall probabilities in directed forgetting, and by investigating the effects of directed forgetting on recognition memory. They utilized the three-list design in order to accurately test the contextual-differentiation theory and to generate data which could be accounted for in a formal model of contextual-differentiation. In addition to addressing issues of recency, their design also addressed another challenge for the contextual-differentiation model: the often observed null effect for recognition memory. Most models assume that context plays an important role in episodic recognition. The assumption is supported by findings that show context-dependent recognition performance (Dennis & Humphreys, 2001; Light & Carter-Sobell, 1970; Murnane et al., 1999). Thus, the context-differentiation model predicted that there should be an effect of the instruction to forget on recognition memory if recognition depends on the use of mentally reinstated context.

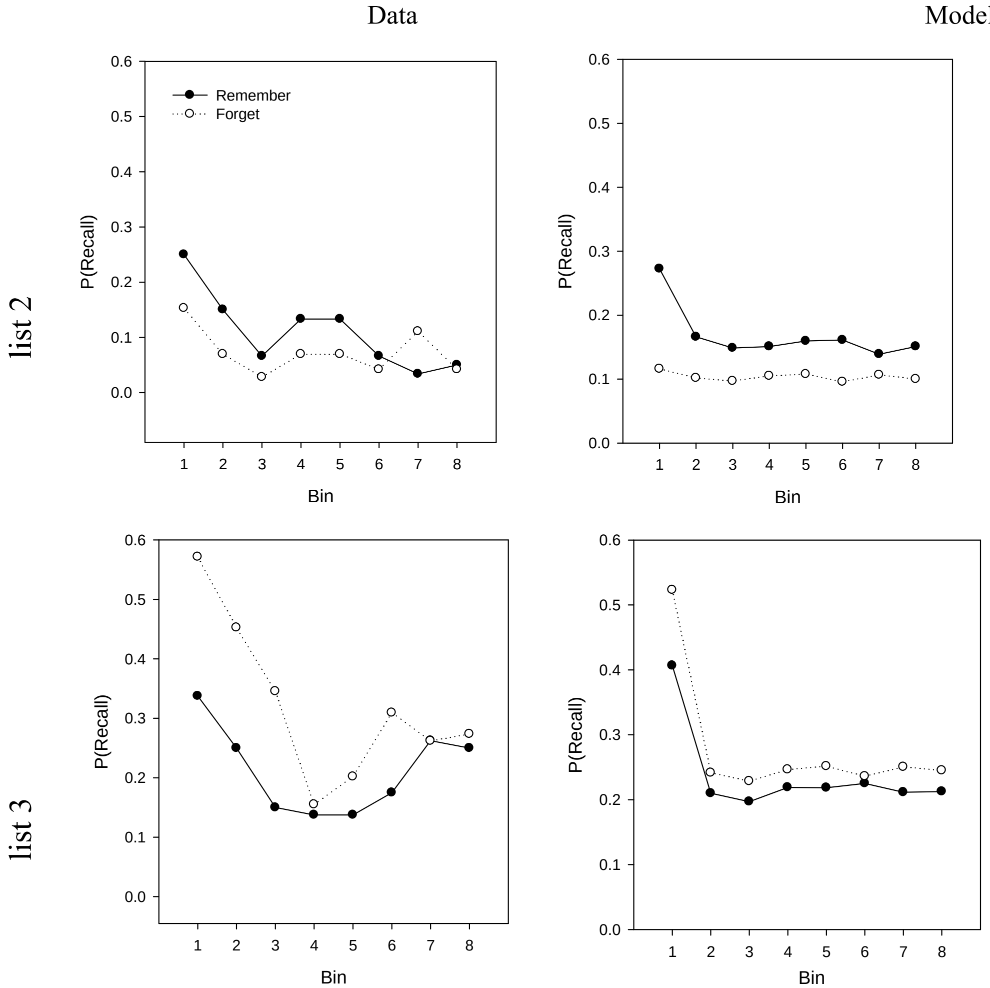
The nature of the recognition tests used to assess directed forgetting is a critical issue. The list method requires multiple study lists. Under these conditions, recognition experiments can use either an *inclusion* test or an *exclusion* test (Jacoby, 1991; Winograd, 1968). In an inclusion test, one should endorse any item studied during the experiment. Hence, context cues that differentiate the study lists are not required. In contrast, exclusion recognition requires the subject endorse only words from a specified list. In this case, the subject must use a context cue that differentiates the study lists in order to accurately perform the task. Note that list-method free recall also requires a context cue for a particular list. Thus, the exclusion task is more similar to what is required for free recall than the inclusion task, and if the contextual-differentiation hypothesis is accurate, then we should see robust effects of directed forgetting on exclusion task performance. Interestingly, most of the recognition experiments in the directed forgetting literature used an inclusion procedure rather than an exclusion procedure.

Lehman and Malmberg (2009) predicted that the effects in an exclusion task should be similar to those observed for free recall, where intrusion rates are reduced by the forget instruction (Malmberg et al., 2006). Thus, there should be costs and benefits on hit rates and the recency advantage for L_2. There should also be more L_2 false alarms when a subject is attempting to recognize from L_1 than there will be L_1 false alarms when a subject is attempting to recognize from L_2, and false alarm rates should be lower in the forget condition.

Lehman and Malmberg (2009) completed serial position analyses on the data collected by Malmberg et al. (2006), including analyses of first recall probabilities. As shown in Figures 3 and 4, these analyses revealed that most of the effect of the forget instruction in free recall is driven by what happens during the initial memory probe. When recalling from L_1, participants in the remember condition are more likely to begin recall by first outputting the first item on the list; the opposite is true when recalling from L_2.

Lehman and Malmberg also examined directed forgetting in a recognition task. Standard effects of directed forgetting were found in an exclusion task: costs and benefits, reflected in decreased hit rates on L_1 and increased hit rates on L_2 for

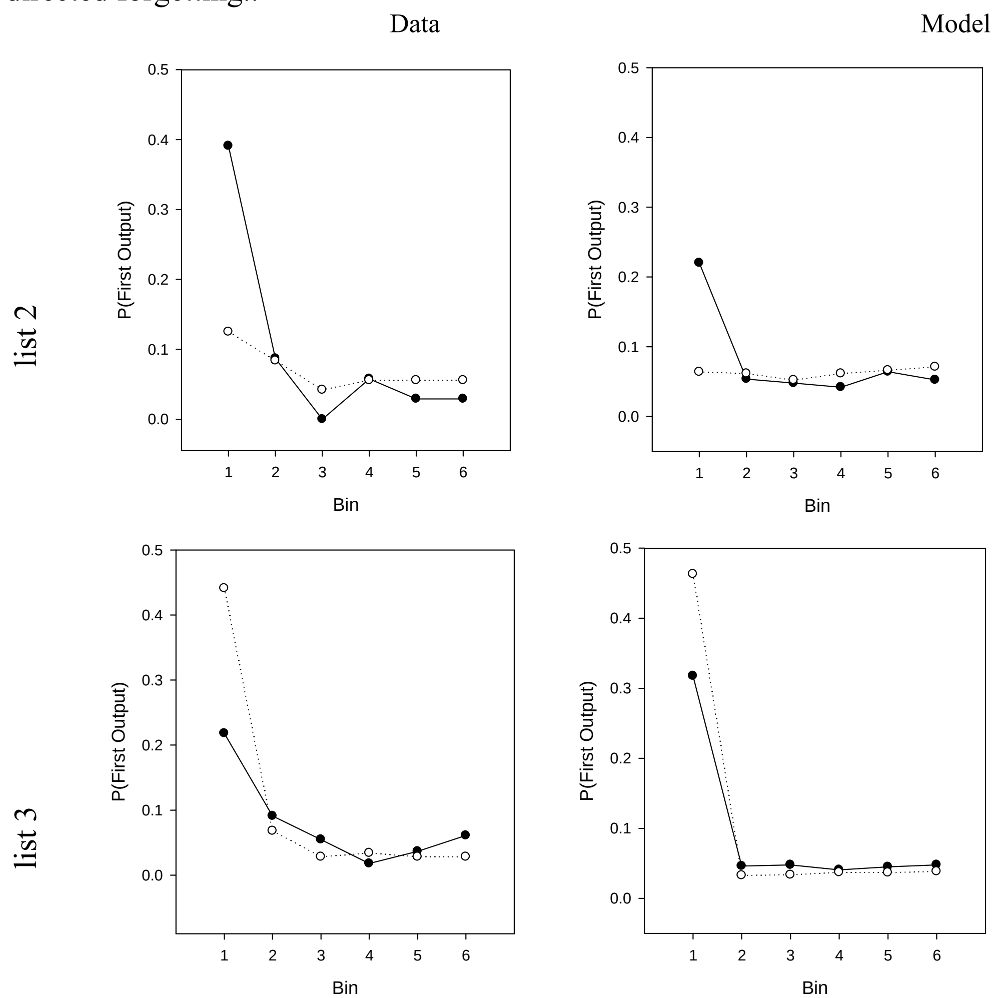
Figure 3. REM model predictions and serial position data in a directed forgetting free recall task.



Note. For the sake of clarity, the 16 item list was compiled into 8 bins spanning two serial positions. For instance, bin *n* contains the data from serial positions *2n-1 and 2n.*

participants in the forget condition. These are shown in Figure 5. Additionally, false alarm rates mirrored intrusion rates seen in free recall: false alarms from L_2 when attempting to recognize items from L_1 were greater than false alarms from L_1 when attempting to recognize items from L_2, and false alarm rates were greater in the remember condition than in the forget condition. As shown in Figure 6, serial position effects in the recognition task did not show the same patterns as in free recall; the directed forgetting effect was not driven by differences only in the beginning of the list.

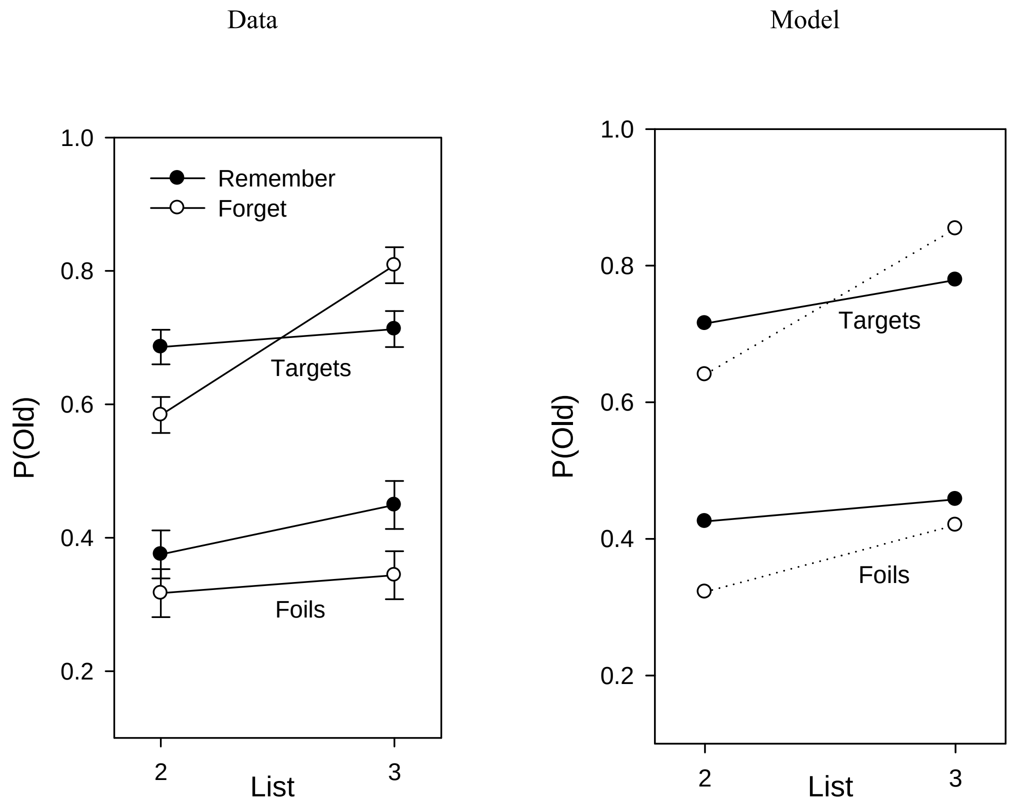
Figure 4. REM model predictions for first recall probability data for free recall in directed forgetting..



Note. For the sake of clarity, the 16 item list was compiled into bins. For first item output, *bin 1* represents the first item on the list (since this is where differences are seen) and all other serial positions are grouped by three.

While Malmberg et al. (2006) developed a model of the retrieval mechanisms supporting the contextual differentiation model for free recall, which was able to account for the costs and benefits of directed forgetting in the contextual differentiation framework, this simple model was not suitable for examining serial position analyses and other fine-grained aspects of the intentional forgetting, or for simultaneously fitting recognition findings. Lehman and Malmberg (2009) developed a context differentiation

Figure 5. REM model predictions for exclusion recognition in directed forgetting.



model of intentional and unintentional forgetting in the Retrieving Effectively from Memory (REM) framework (Shiffrin & Steyvers, 1997). While SAM and REM are both descendents of the A-S model, the SAM model is a mathematical model with provides average recall predictions. The REM model, on the other hand, is a probabilistic computation model which allows us to generate predictions for serial position effects in addition to overall recall probabilities. Additionally, as it was designed to account for recognition findings, which have troubled other models (Shiffrin et al., 1990), the REM model is more suitable for making predictions regarding recognition tasks in directed forgetting. The Lehman-Malmberg REM model was able to account for not only the costs and benefits of directed forgetting in free recall, but also serial position and sequential dependency data, in addition to directed forgetting effects in recognition.

Figure 6. REM model predictions for serial position data for exclusion recognition in directed forgetting.

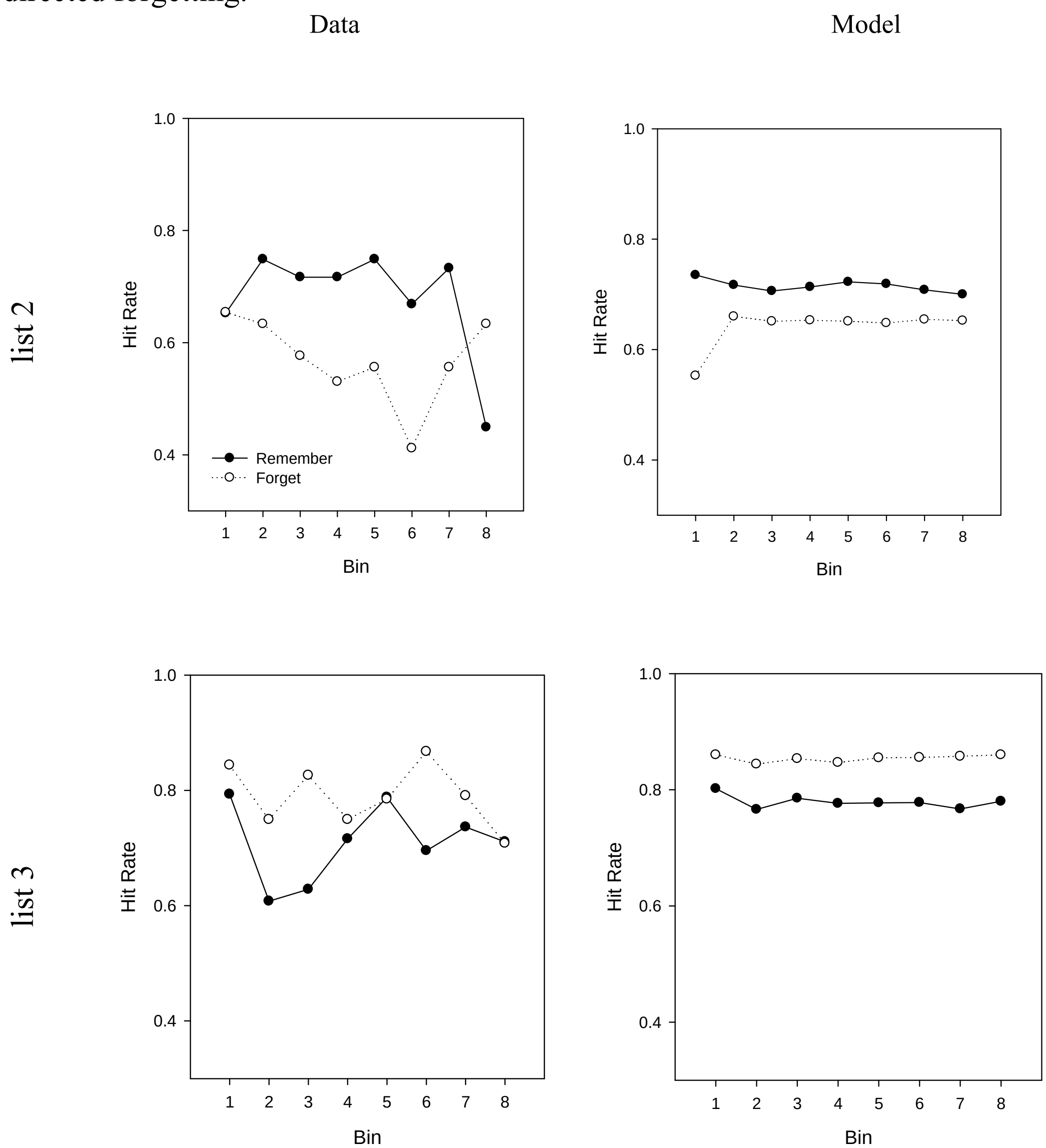
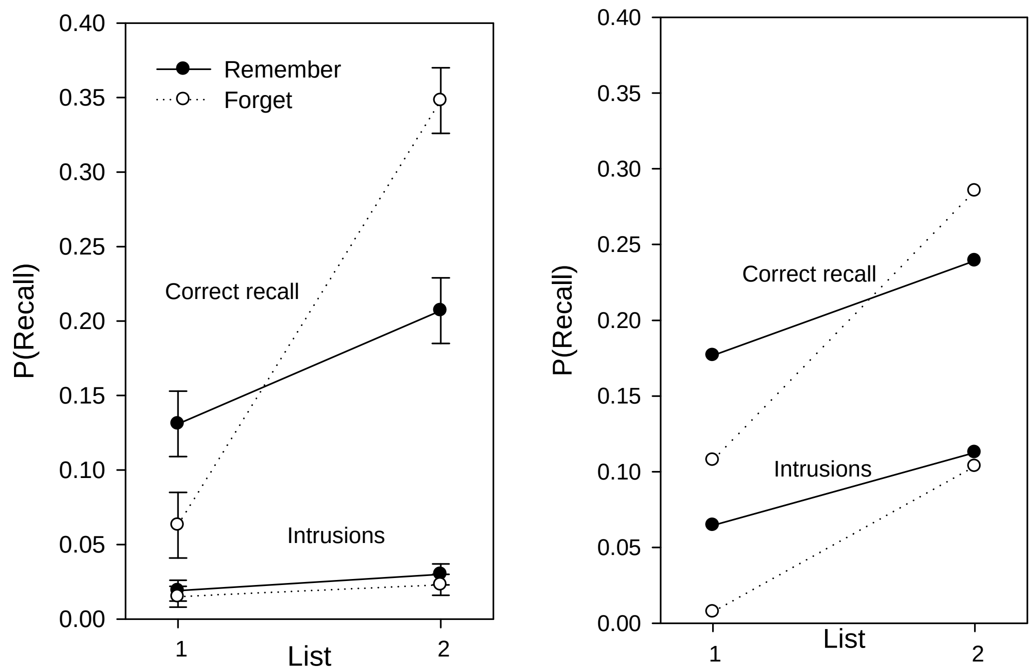


Figure 7. Data and REM model predictions for probability of correct recall and intrusion errors for free recall in directed forgetting.



Note. The intrusions in this graph refer to intrusions that came from either List 1 or List 2. When recalling from List 1, any List 2 item that was output is referred to as an intrusion and vice versa.

1.5 REM Model

1.5.1 Representation

According to REM, general knowledge of items is stored in lexical/semantic memory traces and information about past events is stored in episodic memory traces. Lexical/semantic traces are acquired over a lifetime. They contain information about how words are spelled and pronounced and what they mean. In addition, they contain information about the contexts or situations in which they have been encountered. As such, they are accurate, complete, and generalizable to the contexts in which they usually occur.

These traces are represented by a vector of features. The *w* features comprising the vectors are generated according to a geometric distribution with the base rate parameter, *g*:

$$P[V = j] = (1-g)^{j-1}g,\ j = 1,...,\infty\ .$$

When a word is studied, the *w* item features of its lexical semantic trace are copied to form a new episodic trace that represents this occurrence.

1.5.2 Encoding

Episodic encoding is an incomplete and error-prone process; a feature may be copied correctly, it may be copied incorrectly, or it may fail to be copied. Each lexical/semantic feature associated with an item is copied to an episodic trace with the probability,

$$c\left[1-\left(1-u^{*}\right)^{t}\right],$$

where u^{*} is probability of storing a feature given *t* attempts to do so and *c* is probability of copying that feature correctly. An item will be stored but copied incorrectly from a lexical/semantic trace with a probability *1-c*. If a feature is stored but copied incorrectly, a feature is drawn randomly from the geometric distribution identified by the *g* parameter. If a feature is not encoded, it takes the value zero.

1.5.3 Recognition

A global-matching process is used for recognition memory in REM (Malmberg, 2008; Malmberg, Zeelenberg, & Shiffrin, 2004; Shiffrin & Steyvers, 1997). Whereas sampling of items in free recall is determined by the match between a single item and a given memory cue, recognition judgments are made based on the overall degree of match between all items on the list and the given cue. A decision about whether an item is judged as "old" is made based on the likelihood ratios calculated for all items in the comparison set. The "odds," or the probability that a test item is old divided by the probability that the test item is new, are calculated according to the following equation:

$$\Phi = \frac{1}{n}\sum_{j=1}^{n}\lambda_j ,$$

where λ_j is a likelihood ratio computed for each trace,

$$\lambda_j = (1-c)^{n_{jq}} \prod_{i=1}^{\infty}\left[\frac{c+(1-c)g(1-g)^{i-1}}{g(1-g)^{i-1}}\right]^{n_{ijm}}$$

and where n_{jq} is the number of mismatching features in the j^{th} trace and n_{ijm} is the number of features in the j^{th} trace that match the features in the retrieval cue. Matching features increase and mismatching features decrease the likelihood ratio; cases where no features are stored do not contribute to the likelihood ratio either way. If the odds exceed 1.0, the item is judged as old, otherwise it is judged as new.

1.5.4 Recall

Recall is conceived of a series of sampling and recovering operations in REM (Malmberg & Shiffrin, 2005; Raaijmakers & Shiffrin, 1981; Shiffrin & Steyvers, 1997). Sampling is governed by a Luce choice rule which assumes that the probability of sampling a given trace, *j*, is positive function of the match of trace *j* to the retrieval cue and negative function of the match of other *N*-1 traces to retrieval cue,

$$P(j \mid Q) = \frac{\lambda_j}{\sum_{k=1}^{N} \lambda_k}.$$

Once a trace is sampled, recovery of its contents is attempted (Malmberg & Shiffrin, 2005).

1.6 Lehman-Malmberg Model

1.6.1 Representation

According to the Lehman-Malmberg model, the contextual differentiation in directed forgetting is implemented by integrating context into the model. Memory traces are represented by two concatenated vectors; one vector represents the item, and consists of *w* item features, and the other represents the context in which the item has been encountered, and consists of *w* context features.

When items are studied, context information is stored in episodic traces in the same way as item information. For the sake of simplicity in conducting model simulations, we assumed that context features change between lists with a probability of β, but not within lists (Criss & Shiffrin, 2004; Malmberg & Shiffrin, 2005), although it is likely that context changes slightly within lists (later versions of this model include context that changes gradually within lists; see below). Thus, for each list a single context vector was generated to represent the current context, and all items within that list were associated with the same context information, which is stored according to the rules for item storage outlined above. When a context feature value is changed it is randomly sampled from the geometric distribution. We further assumed that context features change after the final study list, in the same manner as they change between lists.

1.6.2 Encoding

The Lehman-Malmberg model assumes that the content of a stored trace is determined by the operations of a limited capacity buffer (Raaijmakers & Shiffrin, 1981; also Atkinson & Shiffrin, 1968; Malmberg & Shiffrin, 2005). As a descendent of the Atkinson and Shiffrin theory (1968), the interaction of control processes and structural aspects of memory are used to model serial position data. Control processes operate on items located in a limited capacity rehearsal buffer during encoding. The capacity of the buffer is not known, but we assumed for the current purposes that it is two items (see also Atkinson & Shiffrin, 1968; Lehman & Malmberg, 2009). While study items are attended to, they reside in the buffer, and information is encoded about them in one or more episodic traces. Thus, upon the presentation of the first list item, it enters the buffer, and an episodic trace is stored. Assuming that no items repeat, each lexical/semantic feature associated with the first list item and each context feature is copied to an episodic trace according to the equation above, where u_i^* is probability of storing an item feature, and u_c^* is probability of storing a context feature.

Upon the presentation of the second list item, it enters the rehearsal buffer, and a new episodic trace is stored. The trace consists of item information associated with the second list item and context information stored according to the equation above. We further assumed that a result of the capacity limitation is that encoding is split between the storage of item, context, and associative information (Lehman & Malmberg, 2009). In this example, the two buffered items compete for encoding resources. Some of the resources are spent encoding the second list item, and we assume that the resources spent encoding it are similar to those spent encoding the first list item when it was initially presented. The remainder of the encoding resources is divvied up between the storage of associative information and context. This is accomplished in the model by reducing the u_x^* parameter for context features such that $u_c^* < u_{c1}^*$, where the latter term is the probability of encoding a context feature for the first list item, and the former is the probability of encoding a context feature for all other list items. In addition, some of the buffer capacity is spent encoding associative information representing the fact that the first and second items were corehearsed. This is represented by appending to the trace representing the second list item a relatively weak encoding of the first item's lexical/semantic features. Again, this is implemented by reducing the u_x^* value for associative information, u_a^*. With the presentation of the third list item, the oldest item in the buffer is knocked out with probability δ, and the encoding cycle begins anew.

1.6.3 Retrieval

Lehman and Malmberg (2009) extended the REM model to account for directed forgetting in both recall and recognition. The first step of the retrieval process is similar across all test conditions (recall, recognition-inclusion, and recognition-exclusion). A relevant subset of memory is created that consists of the items with the strongest association to the context used as the initial retrieval cue (cf. Shiffrin & Steyvers, 1997; REM.5). In order to create the relevant subset, the current context cue is matched against the context stored in the episodic images. Likelihood ratios are calculated according to the above equation, and those with higher likelihood ratios are most likely to become part of the subset.

1.6.4 Free Recall

The free recall task begins with the creation of the cue with which to probe memory. The initial cue consists of only context features; it is a combination of the current test context and reinstated list context. The proportion of reinstated list context features is represented by the γ parameter. The remaining context features in the cue are from the test context.

Free recall operates in REM cycles of sampling and recovery (Malmberg & Shiffrin, 2005). The initial context cue is matched against all traces in the activated subset in an attempt to sample an item from the given list, and an item is sampled. Lehman and Malmberg (2009) assumed for simplicity that all sampled traces are recovered successfully (however modifications to the recovery process were later made, Lehman & Malmberg, 2011). Thus, when an item is sampled and recovered, and it comes from an incorrect list, the subject undertakes a monitoring process to determine whether it is an intrusion. We assumed that items from the correct list are rarely withheld, and hence if an item is sampled and it is from the correct list, it is output with a probability of 1.0. The probability, η, of making an intrusion error given that an item from the incorrect list is sampled and recovered is a positive function of the overlap in context between lists (represented by this parameter). Thus, η is greater in the remember condition than in the forget condition.

If an item is output, the next cue used to probe memory will consist of both context and the recovered item information. Again, the context portion of the cue consists of both current context features and context features associated with the given list. The item portion of the cue consists of the item vector from the last item recalled. Thus, it is most likely that co-rehearsed items, which share the current item's information, will be sampled next. If no item is output, then the original context cue is used for the next probe of memory. The sample-and-recovery process repeats κ times (Davelaar et al., 2005).

1.6.5 Recognition – Exclusion

For the exclusion task, a subject positively endorses only items that came from a given study list. A global matching process is first used to create the relevant subset of items, using the same context cue that was used to create this subset in free recall. After this set of traces is created, a retrieval cue consisting of only the item information that represents the test word is used to probe memory, and the odds are calculated. This is followed by a monitoring task, as in free recall: after an item is identified as old, an output decision is made in the same manner as was used for free recall. That is, it is dependent on the overlap in context between the two lists, and this is captured by the η parameter at test. This is essentially a recall-to-reject process (Dosher, 1984; Humphreys, 1976; Malmberg, 2008). For the sake of simplicity, however, we did not implement the sampling and recovery processes for the exclusion task, since all of the water is carried by the overlap between the contexts: A large overlap in context means that it is harder to distinguish between the two lists and the false alarm rate will be increased. A description of these processes is found elsewhere (Malmberg, 2008).

1.6.6 Directed Forgetting

The context differentiation model assumes that directed forgetting instructions lead to increased context change between lists and better encoding for L_2 in the forget condition (Lehman & Malmberg, 2009; Sahakyan & Delaney, 2003). As such, the directed forgetting instructions have effects on both encoding and retrieval operations in the model. The context differentiation occurs by an increased rate of context change between lists after the forget instruction, represented by an increased β parameter. Additionally, the encoding of context associated with the first item on a list is increased for the first item on L_2, represented by an increased u^*_{c1}, under the assumption that all other items have been dropped from the buffer. Finally, the forget instruction decreases γ, the probability of reinstating context features used in the cue to probe memory for L_1.

1.6.7 Model Evaluation

The major modeling challenge was to simultaneously account for unintentional and intentional forgetting in a comprehensive and detailed manner. This was difficult because despite the costs and benefits for free recall and recognition, differences remain. For instance, intrusion rates are low for free recall but false-alarm rates are relatively high for exclusion recognition, and the tasks produce different serial position curves. Another challenge was to model the first recall probability functions in a manner that made list discrimination possible and produced costs and benefits. The first recall probabilities are critically important because most of the directed forgetting effect in free recall appears to be driven by them.

Our approach was to account for both tasks with a single contextually driven mechanism. Hence, we refer to this as a "global model" because we are explaining these findings and the relationship between intentional and unintentional forgetting with just a few assumptions. The only differences between the models of free recall and recognition are the assumptions concerning retrieval, and they accounted for a wide variety of episodic memory phenomena (Criss & Shiffrin, 2004; Malmberg, 2008; Malmberg, Holden, & Shffrin, 2004; Malmberg & Murnane, 2002; Malmberg & Shiffrin, 2005; Malmberg & Xu, 2007; Malmberg et al., 2004; Shiffrin & Steyvers, 1997, 1998).

Modeling was accomplished with the use of 16 parameters, 12 of which were fixed in all experimental conditions. Without exception, these scaling parameters are the same or almost same as those used to fit other REM models to data (Criss & Shiffrin, 2004; Malmberg, 2008; Malmberg, Holden, & Shffrin, 2004; Malmberg & Murnane, 2002; Malmberg & Shiffrin, 2005; Malmberg & Xu, 2007; Malmberg et al., 2004; Shiffrin & Steyvers, 1997, 1998). The model predictions for free recall are shown in Figure 7, and model predictions for serial position effects and first recall probabilities, along with exclusion recognition are shown in Figures 3, 4, 5, and 6.

Lehman and Malmberg fit over 250 data points, with only four parameters allowed to vary between the remember and forget conditions in accordance with the assumptions of the model: u^*_{c1}, β, ρ_1, and η. With these parameters, a set of 1000 Monte Carlo simulations for free recall and recognition. The same set of parameter values were used to generate predictions for all experiments.

The model produced the correct patterns of costs, benefits, and intrusions for free recall, and the observed interaction between serial position and the forget instruction, where the costs and benefits are greatest at earlier serial positions. The model also produced the correct patterns of costs and benefits in the hit rates for exclusion recognition and the correct false alarm rates in the remember and forget conditions.

Further, the model predicts relatively flat serial position curves in recognition, where the effect of the forget instruction is not driven by items at particular serial positions.

1.7 Testing Predictions of the Lehman-Malmberg Model

Lehman and Malmberg (2011) tested predictions of the Lehman-Malmberg model in relation to the effects of specific retrieval cues on memory performance in directed forgetting. Based on the way that memory is probed with context cues in the model, Lehman and Malmberg proposed that other more effective memory cues, such as category cues, may eliminate directed forgetting effects. After generating the model predictions related to directed forgetting in categorized lists, the model's predictions were empirically tested.

1.7.1 Categorized Lists

According to REM, once a trace is sampled, recovery of its contents is attempted (Malmberg & Shiffrin, 2005). Lehman and Malmberg (2009) assumed for simplicity that all sampled items are recovered. However, in order to develop a model of recall from categorized lists, it was necessary to more clearly specify the recovery process.

Since the contents are only a noisy incomplete representation of a study event, the contents of some traces are more likely to be recovered than others. The recovery probability is a positive function of the number of features in the sampled trace that match the retrieval cue, x,

$$\frac{1}{1+e^{-x+b}},$$

where b is scaling parameter (Lehman & Malmberg, 2011).

The traditional assumption made by these models of categorized lists is that recovery is more likely to be successful for traces stored on categorized lists (Raaijmakers, 1979; also see Raaijmakers & Shiffrin, 1981, for a discussion of retrieval from categorized lists). In this case, the categorized list advantage falls right out of the model; it is due to the additional matches obtained from the use of readily available category features in the retrieval cue.

The predictions of the model are consistent with directed forgetting data when study lists consist of randomly related items (Lehman & Malmberg, 2009). An assumption is implemented to take into account the nature of categorized lists. Prior models of retrieval from categorized lists have assumed that category-to-item associations are stored (Raaijmakers & Shiffrin, 1981). Here we assume that for items that belong to a categorized list, w additional category features are appended to the item vector. These features are shared by all members of a category, thus within a list where all items are members of the same category, these features will overlap for all items. These features are encoded in the same way as item features, and the likelihood of storing these features is represented by the u^*_{cat} parameter.

Retrieval depends not only on the nature of the list, but also on the cues used to probe memory. If a list is categorized, and a temporal cue is used to probe memory, we assume that the same initial test cue will be used, consisting of current context features and some reinstated context features. If a category cue is used to probe memory, however, a different initial cue is used, which consists of not only the same context features, but also of the additional category features appended to the cue. Additionally, when an item is recalled, the next cue used to probe memory will consist of context features, item features, and category features that are retrieved from the last recalled item, giving a recall advantage to items from categorized lists. Thus, a recovery advantage will lead to higher recall rates for any categorized list over uncategorized lists. Additionally, due to the use of category features in the initial cue, lists probed with a category cue will incur additional recall advantages over lists probed with a temporal cue alone (the additional advantage will be driven primarily by more successful initial recall attempts; see Lehman & Malmberg, 2009).

With these additional assumptions, the model makes various predictions about what should occur in a directed forgetting task when lists are categorized and different cues are used to probe memory. The model predicts costs of directed forgetting in control conditions, where randomly constructed lists are used. When L_1 is categorized, and a temporal cue is used to probe memory (the L_1-temp condition), the model again predicts costs of directed forgetting. When L_1 is categorized and a category cue is used to probe memory (the L_1-cat condition), an effective category cue is available at test, and the model predicts that the costs of directed forgetting should be disrupted[2].

In comparing L_1 performance to L_2 performance in each of these conditions, recency of L_2 is predicted in the control condition (in that performance on L_2, the most

---

[2] One might expect that the category cue should lead to improved performance in both the remember and forget conditions, rather than in only the forget condition. However, the costs of directed forgetting are the result of an ineffective context cue used to initially probe memory. As context alone is often used as a cue only on the first recall attempt (later attempts also use item information), the costs are captured primarily by first recall probabilities (Lehman & Malmberg, 2009). In the remember condition, the initial temporal cue is effective, thus performance is limited mostly by encoding strength.

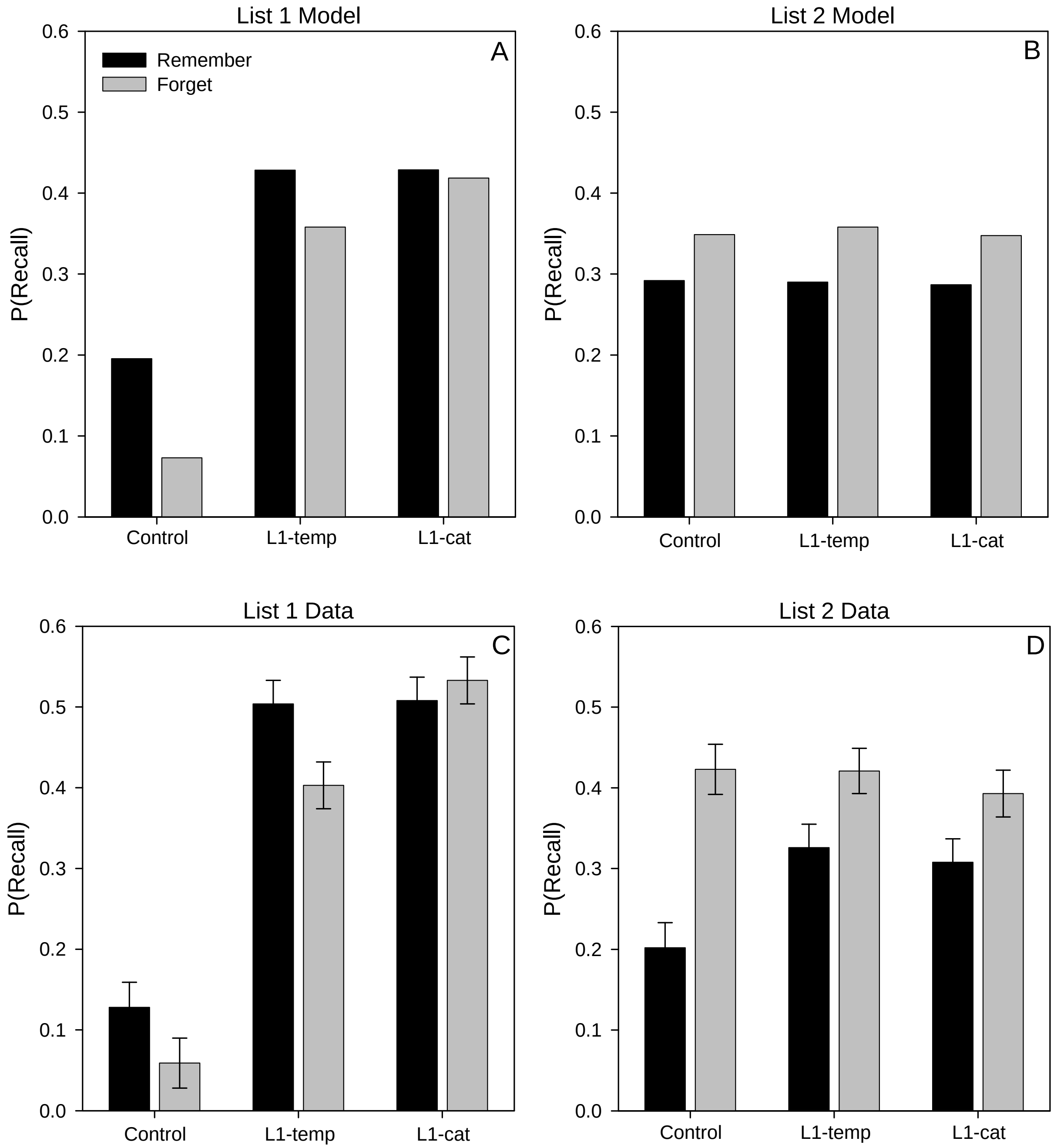
recent list, is greater than performance on L_1, the less recent list; see Lehman and Malmberg, 2009). A recovery advantage for categorized lists should lead to better recall for L_1 in all of the categorized list conditions when compared to L_1 in the control condition.

1.7.2 Empirical Support

We tested these predictions with an experiment in which the lists consisted of either unrelated words or categorical exemplars. Our assumption was that the structured list (consisting of categorical exemplars) would provide additional category cues with which to probe memory. In addition, we varied the instructions given to the subjects at test. In two conditions, the control condition, in which L_1 consisted of randomly related items and in the L_1-temp condition, one of the conditions in which L_1 items were exemplars drawn from a common category (e.g., clothing), subjects were provided a temporal cue at test: Recall as many words from L_1 as you can. In the L_1-cat condition, L_1 was categorized and subjects were provided a category cue at test: Recall as many items from the *clothing* list as you can. The prediction was that the category cue would reduce or eliminate the costs of directed forgetting. These data and model predictions are shown in Figure 8.

We found that categorized lists produced the costs associated with intentional forgetting, but only when memory was cued with temporal context. When category cues were used to probe memory the costs of intentional forgetting were eliminated. Additionally, L_2 was recalled better than L_1 in the control conditions, and L_1 was recalled better in the categorized conditions than in the control condition. The model correctly

Figure 8. REM Model predictions and data from a directed forgetting task with categorized lists.



Note. The top row shows the model predictions for List 1 (Panel A) and List 2 (Panel B) in each cue condition: a control condition (Control), and two conditions where L_1 is categorized and either a temporal cue is given (L1-temp) or a category cue is given (L1-cat) at test. The bottom row shows the data from the experiment. Panel C shows List 1 performance (costs) and Panel D shows List 2 performance (benefits). P(Recall) = Probability of recall. Error bars represent standard error.

predicted the observed patterns of data, and thus proved to be a viable explanation for how intentional forgetting is accomplished and the conditions under which it will and will not occur.

Thus, in addition to accounting for all of the data presented by Lehman and Malmberg (2009), the Lehman-Malmberg model made a priori predictions about what will occur when categorized lists are used in directed forgetting tasks and specific cues were used. The model was able to account for data generated from a directed forgetting task using categorized and uncategorized lists without any additional parameters.

Chapter 2

Evaluating the Lehman Malmberg Model

I will now extend the Lehman-Malmberg model to provide a more comprehensive account of memory processes, and to account for some of the findings that have troubled buffer models in the past. In order to develop the Lehman-Malmberg model used to account for directed forgetting effects, we made a few simplifying assumptions. First, we assumed that context does not change within a list. This assumption was made purely for the purposes of creating a simpler model; however it is not consistent with the general view that context fluctuates over time (Dennis & Humphreys, 2001; Howard & Kahana, 2002; Mensink & Raaijmakers, 1989). The current version of the model assumes that context drifts slowly over time not only between lists, but also within a list, and the rate at which context changes may be increased by the task. For example, context drifts more quickly between lists, and tasks such as directed forgetting will increase context change to a greater degree. Next, we assumed for simplicity that the capacity of the buffer was two items. As the buffer is viewed as a control process that may be used differently depending on the task, a buffer size of two may sometimes be appropriate; however, it is likely that a larger buffer is sometimes needed (Atkinson & Shiffrin, 1968). Thus, the current model includes an increased buffer size of three items and allows the capacity to change according to the demands of the task.[3] Finally, while the Lehman-Malmberg

[3] One might argue that even a buffer size of three is unrealistic, given Miller's (1956) seven-plus or-minus-two theory, another staple of introductory textbooks. Miller's data, however, shows that one can remember *lists* of five to nine items. Other work suggests that when using longer lists, the number of items that can

model included a buffer component that influenced encoding, the buffer did not play a role in retrieval, as the directed forgetting tasks for which this model was derived involved only delayed free recall. In order to account for immediate free recall, buffer operations in retrieval were added to the model.

The model assumes that in immediate free recall, the contents of the buffer are retrieved in such a way that differs from retrieval of items in long-term storage. Thus, recall is usually initiated by sampling only from the buffer, using only the most recent context as a cue. After sampling from the buffer occurs, retrieval continues as it does in delayed free recall, with context cues that consist of a combination of the current context and reinstated beginning of list context. The effect of delay is represented in the model by the storage of additional traces after the study list, which are generated in the same manner as list items. On the assumption that the distractor task eliminates items from the rehearsal buffer, there is no "dumping" of the buffer as there is in immediate testing. Parameter values are listed in Table 1.

2.1 Craik & Watkins (1973)

Important to tests of buffer models is the distinction between *maintenance* rehearsal and *elaborative* rehearsal (Craik & Lockhart, 1972). Whereas elaborative rehearsal serves to enrich a memory trace, increasing the likelihood of retention, the goal of maintenance rehearsal is to maintain the trace in temporary representation. Craik and Watkins (1973) conducting an experiment that required participants to use maintenance rehearsal in an incidental learning task. Participants were required to study lists of

---

be *simultaneously accommodated* is closer to three or four (Cowan, 2001; Crowder, 1989; Wickelgren, 1964), and this is what is represented by the buffer in our model.

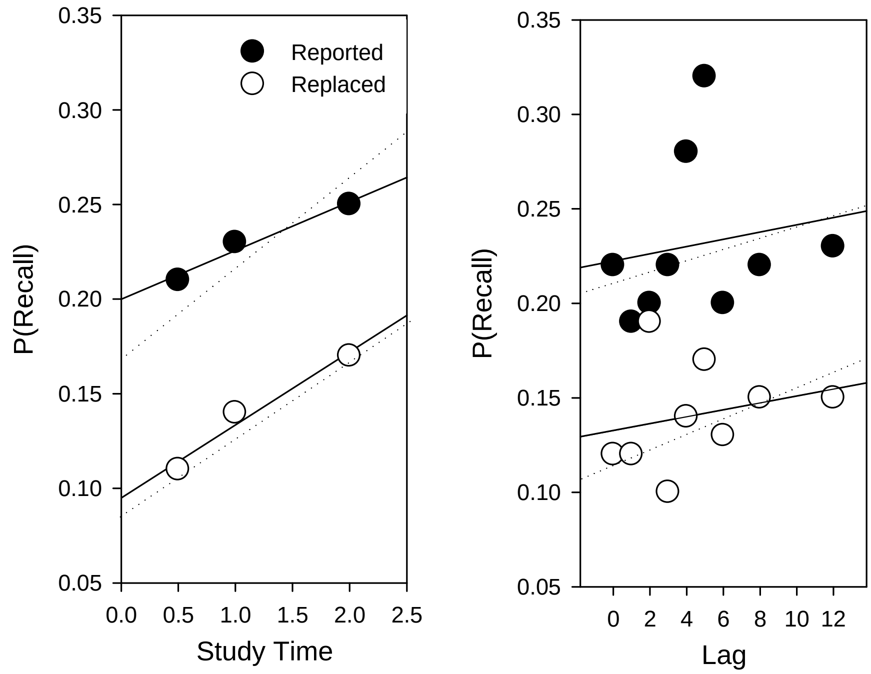
Table 1. Parameter Values and Descriptions for Lehman-Malmberg Model

Parameter	Value	Description
g	.4	Environmental base rate (standard value)
w	8	Number of item and context features
c	.8	Probability of correctly storing a feature
u^*_i	.5	Probability of storing an item feature
u^*_c	.3	Probability of storing a context feature
u^*_a	.1†‡	Probability of copying a co-rehearsed item's feature
u^*_{c1}	.5‡	Probability of storing a context feature for first item on a list
t	3	Number of storage attempts
κ	10	Number of sampling attempts
β_w	.2‡	Probability of change for context features within lists
β_b	.5	Probability of context change between lists, or after a list
γ_1	.4‡	Probability of reinstating context features from beginning of list
γ_m	.4	Probability of reinstating context features from recovered item
b	5	Scaling parameter for recovery

Note. † Parameter values that differ in delay and continuous distractor conditions. For delay, $\gamma_1 = .2$, and 10 additional items are stored after the list. For continuous distractor, $\gamma_1 = .3$. ‡Parameter values that differ in paired-list condition. For pairs, $u^*_a = .5$, $u^*_{c1} = .6$, β_w between pairs = .3. For the Craik and Watkins (1973) simulations, a single u^* value of .1 was used for both item and context information, and 2 of the w item features representing the shared first letter remained the same for all critical items on the list.

words, which included multiple "critical" words which were identified by their first letter. At the beginning of each list, the critical letter was indicated, and participants were instructed to report only the most recent word that began with that letter. This required participants to maintain each critical word until another word with that letter was presented. The number of words which an item was required to be maintained in memory will be referred to as *lag*. Craik and Watkins varied the lag, presumably varying the amount of maintenance rehearsal that was required for each item. Additionally, the words were presented at a rate of one every half second, one every second, or one every two seconds. At the end of the experiment, memory was tested for all of the critical

Figure 9: Data from Craik and Watkins (1973) and model predictions from the Lehman-Malmberg Model



Note: regression lines based on Craik and Watkins' (1973) data are shown as solid lines; model predictions are represented by dotted lines.

words in the experiment, including both those that were reported and those that were replaced and not reported. The data from this experiment is shown in Figure 9.

Craik and Watkins (1973) reported that increases in study time were related to increases in recall for both reported and replaced words, but that lag did not affect recall. Craik and Watkins suggested that these findings were problematic for the A-S model because it predicted that items that spent more time in the buffer should be better remembered. These claims, however, are problematic themselves, as they are somewhat inaccurate. First, as shown by the solid regression lines in Figure 9, recall increases with increasing lag. Next Atkinson and Shiffrin (1968) explicitly stated that information may only be weakly encoded when incidental encoding is used. Thus, the Craik and Watkins (1973) data is consistent with the predictions of the A-S model.

We extended the model to account for the data from Craik and Watkins (1973) by assuming that when subjects are using maintenance rehearsal in an incidental task, less encoding will occur than when using elaborative rehearsal. Because only one word needed to be rehearsed at a time, we assume that the buffer size is one, and that additional item information may be encoded with increased time in the buffer. Thus, we varied the t value, which represents the number of units of storage time for each item, for item and context features, such that increases in study time increased the storage of item features to a greater degree than context features (Malmberg & Shiffrin, 2005). Higher values of t were used for context features in the reported condition than in the replaced condition, assuming that reporting the word at the end of the list increases context storage. Thus, the t value for item and context features was calculated as follows: For item features, $t = lag + 1$. For context features, $t = studytime * a$, where a is greater for reported ($a = 2$) items than for replaced items ($a = 10$).

The model assumes that no information is stored for non-critical items, as these items are never present in the buffer. Additionally, the model assumes that during retrieval, subjects are aware of some of the first letter information that was used during the study lists, and they use this information as part of the cue used to probe memory. As shown in Figure 9, where the dotted line represents the model's predictions, the model provides a good fit to the data, X^2 (15) = 4.06, p >.05.

Thus, the model is able to account for data once thought to be troublesome to buffer models of memory simply by manipulating the way that the buffer is used during maintenance rehearsal. Increases in the *t* value with increased time spent in the buffer occur differently for item information and context information (Malmberg & Shiffrin, 2005), leading to the prediction that increases in lag have small effects on recall when maintenance rehearsal, but reporting items causes additional storage of context information, which has greater effects on recall.

2.2 Continuous Distractor Task

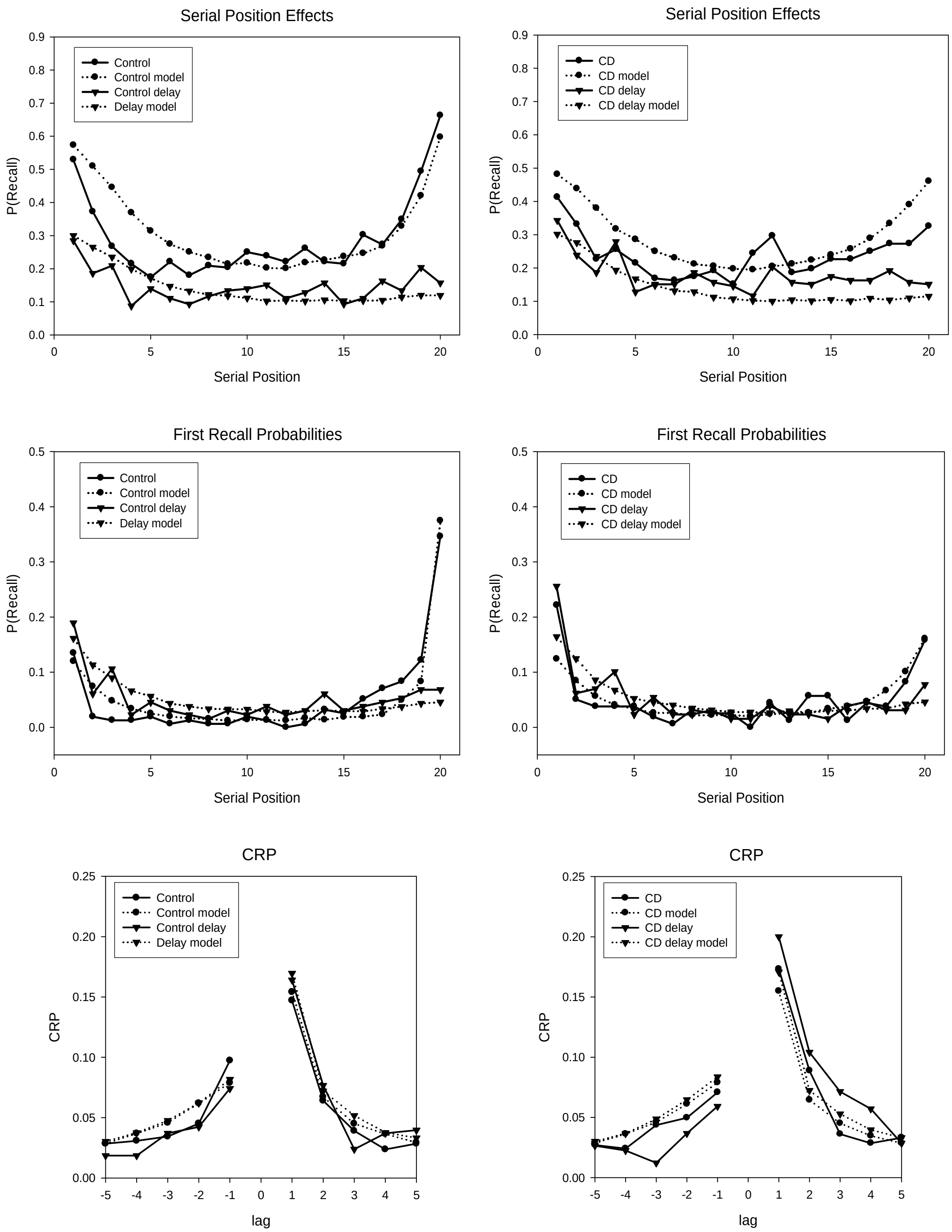
We now wish to further examine the buffer process and address long-term recency and lag-recency effects in the model. Our goal is to address two questions related to the buffer. First, is the buffer used as a control process during encoding, such that encoding resources can be differentially allocated depending on the requirements of the task, or can we account for the data from continuous distractor experiments without such a process? Second, do items in recent memory exist in a short-term buffer – a privileged state such that they are differentially accessible compared to items learned less recently?

To address the first question, we generated model predictions for the continuous distractor paradigm, in which a short distractor task is presented after each item on the list, including the last item. We conducted an experiment comparing serial position effects (SPs), first recall probabilities (FRPs), and CRPs in standard free recall tasks to those in continuous distractor tasks. Findings that have said to be challenging to buffer models include long-term recency effects and contiguity effects in the continuous distractor task. Despite the distractor task presented after each item on a list, which, according to Howard and Kahana (1999) should prevent the encoding of associations, contiguity effects are still seen. Additionally, the presence of the distractor task should eliminate the final item on the list from the buffer, eliminating the recency effect (Bjork & Whitten, 1974; Howard & Kahana, 1999).

The continuous distractor task is represented in the model by reducing the buffer capacity to two items under the assumption that it will be harder to maintain more items in the buffer while completing the distractor task between items. Additionally, for both immediate and delayed free recall in a continuous distractor task, the model assumes that there are no items remaining in the buffer at the end of the last distractor task, thus the first memory probe uses the combined context cue used in the delayed condition (described previously), with a reduced likelihood of reinstating features from the beginning of the list, due to the increased context change that has occurred throughout the list as a result of the continuous distractor task. Model predictions for continuous distractor lists versus control lists in both immediate and delayed free recall are displayed by the dotted lines in Figure 10. In order to test the model's predictions, an experiment was conducted examining the continuous distractor task in both immediate and delayed free recall. This experiment examined both standard and continuous distractor free recall in both immediate and delayed free recall conditions. Rather than a math task, commonly used in such experiments, the distractor task interspersed between items for the continuous distractor conditions required participants to provide rhymes for irrelevant words. We hoped that the verbal nature of this task would make it harder for participants to rehearse the list items during the distractor period.[4] Additionally, a delay condition was included in order to examine the effect of a long distractor period on the recency effects. While a delay of 10 seconds (a time commonly used in the continuous distractor

[4] We thank Doug Nelson for this suggestion.

Figure 10: Continuous distractor data and model predictions



task) may not be sufficient to eliminate items from the buffer (Glanzer & Cunitz, 1966), a longer delay may be more effective. Thus, we chose a 5-minute delay in order to assess this possibility.

2.3 Experiment 1

2.3.1 Method

2.3.1.1 Participants and Materials. Participants were 86 undergraduate psychology students at the University of South Florida who participated in exchange for course credit. For each participant, eight word lists were created, each consisting of 20 randomly related concrete nouns (between 20 and 50 occurrence per million; Francis & Kucera, 1982). Additionally, four lists of rhyme words were created. Each rhyme list consisted of 20 monosyllabic words with a rhyme-set size of at least 12 (Nelson, McEvoy, & Schreiber, 1998). The experiment was presented on a computer in an individual subject booth, and the rhyme lists were printed in paper booklets.

2.3.1.2 Procedure. The four conditions were manipulated within subjects in blocks, where two lists were presented in each condition. The order of the conditions was counterbalanced. At the beginning of the experiment, participants were told that they would be studying multiple lists of words, and the instructions for each list would appear before the list. For all conditions, words appeared on the screen one at a time for 1s. For all conditions, 60s were allowed for recall. After each test, they were given their percentage score for the list and told to try to improve their score for the next list.

In the *immediate control* condition, words appeared on the screen one at a time with a .5s ISI. Immediately after the list was presented, a free recall test was given for that list, in which participants were instructed to enter all of the words they remembered from that list onto the screen.

In the *delayed control* condition, words appeared on the screen one at a time with a .5s ISI. After the list was presented, they completed a 5 minute distractor task. During the task, they watched a 4.5-minute "How It's Made" video, with a 30s quiz afterward (to assure that they were attending to the video). After the distractor task, they completed the same free recall task as described in the *immediate control* condition.

In the *immediate continuous distractor* condition, participants were to alternate between memorizing a word on the screen and writing rhymes for a different word in the printed booklet. Before beginning the continuous distractor condition, participants read explicit instructions detailing the procedure, followed by a quiz to be sure they understood the procedure. To encourage participation, they were told that they needed to reach a certain number of rhyming words in order to complete the experiment. They then saw a demonstration, and completed a two word practice list, after which the experimenter checked to be sure that they were attempting to memorize the only words on the screen and write rhymes for only the words in the booklet during the practice trial. Once they correctly completed the task, they began the study list. Words appeared on the screen one at a time, with a 10s delay after each word. During this delay, the ***** symbol appeared on the screen alerting participants that they should now turn to their rhyme booklets and begin creating rhymes for the next word in the booklet. After 10s, a tone alerted them to look back at the screen for the next word. This repeated throughout the list, so that after each studied word, they had to provide rhyming words for a word in the booklet (including after the last item on the list). Participants then completed the same free recall task as described in the *immediate control* condition.

In the *delayed continuous distractor* condition, the procedure was the same as in the *immediate continuous distractor* condition, except that they completed the same 5-minute distractor task after study as in the *delayed control* condition, followed by the same free recall task described above.

2.3.2 Results

In the control conditions, recall was significantly greater in the immediate condition than in the delayed condition, $t(1,85) = 11.445$, $SD = .12$, $p < .001$. Serial position analyses revealed a serial position x condition interaction, $F(28,2380) = 5.06$, $MSE = .194$ $p < .001$. As shown in Figure 10, both primacy and recency are present in the immediate testing condition. In the delayed condition, the recency effect was eliminated. All differences were significant at alpha = .05. A significant serial position x condition interaction is also present in first recall probabilities, $F(28,2380) = 8.15$, $MSE = .046$ $p < .001$.. Figure 10 shows that in the immediate condition, participants were most likely to initiate recall with the last item on the list, whereas in the delayed condition, they were more likely to begin recall at the beginning of the list. For conditional recall probabilities, there was not a significant lag by condition interaction, $F(37,3145) = 1.59$, $MSE = .024$, $p = .11$, as shown in Figure 10. Because of the marginal p value, planned comparisons were conducted, revealing no significant differences between the two conditions at any lag.

In the continuous distractor conditions, recall was significantly greater in the immediate condition than in the delayed condition, $t(1,85) = 4.604$, $SD = .12$, $p < .001$.

Serial position analyses revealed no significant serial position x condition interaction, $F(28,2380) = 1.45$, $MSE = .145$ $p = .10$. While the interaction was not significant, planned comparisons revealed differences in recency (the last item on the list.), but no other differences were significant, as shown in Figure 10. A significant serial position x condition interaction is also present in first recall probabilities, $F(28,2380) = 1.68$, $MSE = .046$ $p = .03$. Again, planned comparisons revealed differences in first recall probability for the last item on the list, but for no other serial positions. For conditional recall probabilities, there was not a significant lag by condition interaction, $F(37,3145) = 1.76$, $MSE = .025$, $p = .07$, as shown in Figure 10. Because of the marginal p value, planned comparisons were conducted, revealing no significant differences between the two conditions at any lag.

2.3.4 Discussion

The model was able to capture all of the effects in a continuous distractor task, including the lag recency effect and the long-term recency effect (in fact, the model predicts greater long-term recency than is actually present in the data). The key assumption that allows the model to account for these data is that the buffer size is reduced during the continuous distractor task, but that associations between items are still possible. The model can simultaneously account for two sets of findings which have been said to be problematic for the buffer model, with a few minor changes to prior versions of the model aimed at developing a more realistic account of memory processes. Proponents of single-store models have suggested that long-term recency effects in a continuous distractor paradigm are troubling for dual-store models, as such models assume that recency effects are due to the presence of items in a short-term buffer at the time of test (Crowder, 1989; Howard & Kahana, 1999). Crowder (1989) explicitly states, "The traditional association of the recency effect in free recall with some transient memory has now been discredited by the work of Bjork and Whitten (1974)" (p. 274).

We have shown here that these findings are not troubling for a model that assumes that the buffer is still utilized in the continuous distractor task. Koppenaal and Glanzer (1990) showed that changing the distractor task after the last item on a list eliminates long-term recency effects. They hypothesized that after repeated exposure to a distractor task, subjects habituate and become able to simultaneously rehearse items in the buffer and complete the distractor task. Thus, they suggested that recency is due to retrieval from a temporary rehearsal buffer even in the long-term recency paradigm. Our findings are consistent with such a proposal. First, long-term recency effects were eliminated when a sufficiently long distractor task (5 minutes) was used. Additionally, in a questionnaire given to participants after they completed the continuous distractor task in pilot work, over half of participants reported using the time when they were working on the interspersed math problems to rehearse items from the list. This was our motivation for using a rhyme task in the experiment reported in this manuscript; even with the use of the rhyme task, designed to discourage rehearsal, many subjects still reported trying to maintain the to-be-remembered words while completing each rhyme task, so that they could make associations with other words on the list. For example, when asked about strategies used, one subject reported, "I tried to keep the memorize word in my head while I came up with rhymes for the rhyme word so that I could connect it to the next memorize word when it came up." Thus, while the rhyme task was more effective than a math task at preventing rehearsals, subjects were still able to make use of the buffer during this task.

Regardless, we do not wish to argue that long-term recency effects are due to some transient memory phenomenon that would be eliminated with a sufficiently distracting task; rather, we argue that long-term recency effects do not preclude the existence of some transient memory phenomenon. As suggested by Cowan (1995), even though two memory phenomena may be made to mimic each other, differences in the mechanisms that elicit these phenomena may exist. Although recency effects in immediate free recall and continuous-distractor free recall may be similar in appearance, this does not necessarily indicate that they are due to the same mechanism. We argue that recency in immediate free recall is due to retrieval from a rehearsal buffer and long-term recency in continuous-distractor free recall is a long-term memory phenomenon. While both lead to increased memory for the final items on the list, they arise from different processes. Similarities in recency effects for immediate free recall and continuous distractor free recall may be taken as evidence against a dual-store model only if they display the same properties (Cowan, 1995). However, there are various sources of evidence suggesting that recency effects in these two tasks do not display the same properties.

First, immediate recency and long-term recency are differentially affected by the presentation of an irrelevant auditory stimulus at the end of an auditorily presented list (Glenberg, 1984). Whereas recency effects are reduced for immediate recall when the same irrelevant stimulus is presented for each trial, recency effects are only reduced in continuous distractor recall when a different stimulus is presented after each trial. Next, the effects of output order on recency effects differ for immediate and continuous distractor tasks. In immediate tasks, a recency advantage is obtained when subjects are required to initiate recall with items from the end of the list rather than the beginning of the list (Dalezman, 1976). In continuous distractor tasks, however, no recency advantage is present when subjects are required to initiate recall with the end of the list versus the beginning of the list (Whitten, 1978), suggesting that recency effects in immediate recall are subject to output interference, but this is not the case for recency effects in continuous distractor recall. These findings are consistent with the viewpoint that immediate recency effects are due to the presence of the last few items in a short-term buffer at time of test, whereas long-term recency effects arise from a different process. Finally, other work has shown that word length differentially affects immediate recall and continuous distractor recall (Cowan, Wood, & Borne, 1994), and as does semantic relatedness between words on a list (Davelaar, Haarman, Goshen-Gottstein, & Usher, 2006).

Thus, much empirical evidence suggests that immediate recency and long-term recency are produced via different processes. In fact, Bjork and Whitten (1974) stated that long-term recency effects and immediate recency effects “reflect entirely different memory processes” (p. 183). Additionally, the Lehman-Malmberg model produces both immediate recency effects and long-term recency effects through two different mechanisms that correspond to those used to describe differential recency effects in these two procedures; immediate recency effects are related to retrieval from the buffer and long-term recency effects are due to contextual similarity between the last items on a list and the retrieval cue. It is apparent in Figure 10 that while a recency effect was present in the continuous distractor task with immediate testing, the magnitude of this effect was notably smaller than in the control condition (see also Howard & Kahana, 1999), suggesting that different explanations for these two effects are warranted.

As a further test of the model, we examined its prediction that encoding resources can be differentially allocated depending on the requirements of the task (Atkinson & Shiffrin, 1968). We used single-item study lists and paired-item study lists in order to encourage different rehearsal strategies. If memory utilizes control process in the form of a rehearsal buffer, then items may be both encoded and recalled differently depending on the nature of the list.

2.4 Single Versus Paired Item Study Lists

The critical difference between studying a list of single items and a list of paired items in the model is the way the buffer operates. For single items, the buffer capacity is three items. For paired items, the buffer capacity is two items, and two items within a pair always share the buffer. Earlier versions of the Lehman-Malmberg model did not include chunking operations, but it is now necessary to explore how chunking functions in the model. Many features of classic chunking theories already exist in the model, though the nomenclature may differ. Johnson (1970) describes chunks as "items or information sets which are stored within the same memory code" (p. 172), where a memory code is a mental representation of information that was learned, analogous to a *trace* in our model. As with memory traces in the REM model, codes are representations of information which are distinct from the information itself (lexical/semantic traces in REM can include errors or missing information). Johnson also describes *recoding*, the process of learning a code for a chunk, and *decoding*, the process of translating the code into the information it represents, which roughly correspond to the encoding and retrieval processes in REM.

We implement some new assumptions from original theories about chunking into this model. The two main assumptions, adapted from Johnson (1970) and the way they are implemented are discussed below.

Assumption 1: Associations are made between items in the same chunk. If items are from different chunks, the associations between them are minimal.

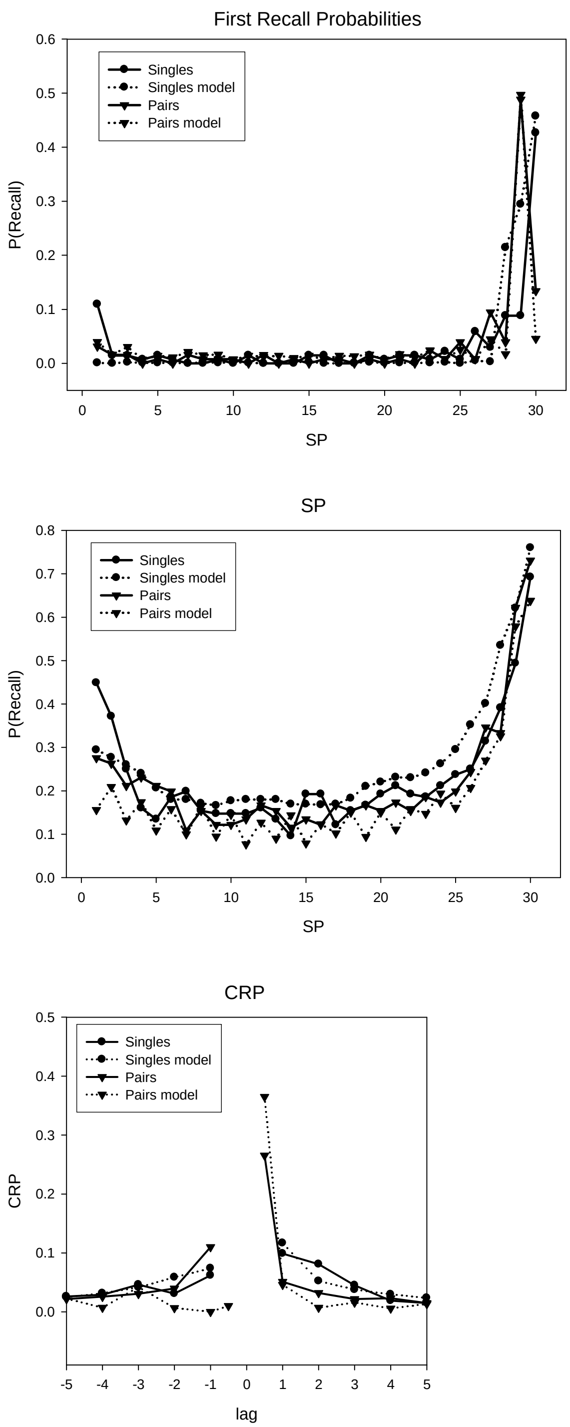
Implementation: During encoding, both members of a pair are stored in a single trace. Due to the staggered presentation, encoding occurs as follows: for the first item in a pair, item information and item-context associative information is stored. When the second item in a pair enters the buffer, item information and associative information from the first item in the pair is stored. Thus, the context is more strongly associated with the first item in a pair than the second item in the pair, and the first item in the pair is associated with the second item, but for simplicity, we assume in the current model that the second item in a pair is not associated with context. Accordingly, $u*_a$ for pairs > $u*_a$ for single items, and $u*_c$ for the first item in a pair > $u*_c$ for the second item in a pair (which is currently set to zero).

Assumption 2: Chunks are recalled in an all-or none manner. In order to recall any information from within a chunk, it is necessary to recover the code from memory. If the code is recovered, then at least implicit recovery of all information represented by the code occurs. Recall from chunks begins with the first item; the other items are maintained in short-term memory while the first item is being retrieved.

Implementation: During retrieval, the current context is initially used as a cue. When an item that is the first member of a pair is successfully retrieved, the next item to be sampled is the second item from that pair; the recovery process occurs as it does in the single-item model.

The model uses most of the same parameter values for pairs as for single items, aside from those listed above, which vary with for paired items (parameter values are listed in Table 1). The model predictions are shown as dotted lines in Figure 11. The model makes two notable predictions. First, as shown in the top graph, FRPs differ for single and paired items. For single items, the item on the list that is most likely to be output first is the most recently studied item. However, for paired items, the item that is most likely to be output first is the penultimate item, or the first item from the last pair. This prediction is derived from Assumption 1 described above. Next, the model predicts that for paired items, are much more likely to make a +1 lag transition (representing transitioning within a pair) than in single items. In order to test the predictions of this model, we conducted an experiment using single and paired item study lists.

Figure 11: Single versus pairs data and model predictions



2.5 Experiment 2

2.5.1 Method

2.5.1.1 Participants and Materials. Participants were 39 undergraduate psychology students at the University of South Florida who participated in exchange for course credit. For each participant, eight word lists were created, each consisting of 30 randomly related concrete nouns (between 20 and 50 occurrence per million; Francis and Kucera, 1982). The entire experiment was presented on a computer in an individual subject booth.

2.5.1.2 Procedure. At the beginning of the experiment, participants were told that they would be studying multiple lists of words. All participants studied four lists of single words in one block and four lists of paired words in another block. The blocks were counterbalanced, with the instructions at the beginning of the block. Instructions for the single lists informed participants that they would be shown the words one at a time, and they should try to create a sentence in order to memorize each word. Instructions for the paired lists informed participants that they would be seeing pairs of words, and they should try to create a sentence in order to memorize both words.[5]

For the single word lists, the words appeared on the screen one at a time. For half of the lists, the words remained on the screen for 1s with an ISI of 375ms. For the other half of the lists, the words remained on the screen for 875ms with an ISI of 500ms. This was done so that half of the lists would have an equal study time to that of the words in the paired lists, and the other half would have an equal total study list time to that of the

[5] These instructions were chosen in order to try to standardize strategies used by participants. To be sure that the results from this study were not an artifact of the instructions, another experiment was conducted utilizing the same procedures, but with no instructions given to participants regarding how to memorize the words. No differences were present in the results of the two experiments.

words in the paired lists. As there were no differences in the results between these different study times, these data were collapsed across study times and were not further analyzed.

For the paired word lists, the words appeared on the screen in a staggered fashion in order to maintain a temporal order to the words. The first word in a pair appeared on the screen and remained. After 250ms, the second word of the pair appeared on the screen adjacent to the first word. After 1.75s, the first word disappeared from the screen, so that the second word from the pair remained alone on the screen. After 250ms, the second word also disappeared from the screen. After an ISI of 500ms, the process continued for the next pair. This staggered presentation was used so that the words would appear in pairs but maintain a temporal order like that of the single word lists.

Immediately after each list was presented, a free recall test was given for that list, in which participants were instructed to enter all of the words they remembered from that list onto the screen. They were given 60 seconds to do this. After the test, they were given their percentage score for the list and told to try to improve their score for the next list. They then completed a 30s math task before beginning the next list.

2.5.2 Results and Discussion

There was not a significant difference in the proportion of words recalled from single or paired word lists, $F(1,38) = 2.49$, $MSE = .037$, $p = .123$. There was a significant effect of serial position, $F(29,1102) = 25.02$, $p < .001$, but there was no significant serial position by condition (single or paired list) interaction, $F(29,1102) = 1.33$, $p = .117$.

Figure 11 shows a significant primacy and recency effect for both single and paired list conditions.

For first recall probabilities, there was a significant effect of serial position, $F(29,1102) = 37.16$, $MSE = .007$, $p < .001$, and a significant serial position by condition interaction, $F(29,1102) = 19.01$, $p < .001$. As shown in Figure 11, participants in the single list condition were most likely to initiate recall with the last item on the list; whereas participants in the paired list condition were most likely to initiate recall with the second to last item on the list (the first item in the last pair).

For conditional recall probabilities, there was a significant effect of lag, $F(59,2242) = 37.92$, $MSE = .001$, $p < .001$, and a significant lag by condition interaction, $F(59,2242) = 8.67$, $p < .001$. Figure 11 shows more +1 transitions in the paired list condition (within-pair transitions) than in the single list condition.

Thus, as predicted by the model, recall patterns differ for single item and paired item study lists, as revealed by SPs, FRPs, and CRPs. The model accurately fits the data for both single item and paired item study lists. As predicted by the model, for single items, the first item output is most likely to be the last item on the list, and for pairs, the first item output is most likely to be the penultimate item on the list (the second item from the last pair). The model predicts a *zigzag* effect in the SP curves (Davelaar et al., 2006). For paired items, we see an up and down pattern throughout the list, where the first item in a pair is more likely to be recalled than the second item in a pair. However, due to the small number of participants used in this experiment, there appears to be too much noise in the data to detect these zigzag patterns. Like the model, the data shows a greater likelihood of making a +1 transition for pairs than for single items. Thus, the model is capturing all of the features that distinguish recall patterns in single item lists from those of paired item lists. At this point, the model has been shown to account for a variety of directed forgetting data, including data generated to confirm a priori predictions of the model, in addition to data from paradigms that have been said to be troublesome for buffer models, including tasks manipulating maintenance rehearsal, and the continuous distractor task. Finally, the model has been extended to account for findings in a novel paradigm, with the implementation of assumptions regarding chunking process in memory.

In order to increase our confidence in the Lehman-Malmberg model, it is useful to test a priori predictions of the model in relation to the buffer component, and to further explore the contributions of the buffer to the retrieval process. Specifically, we would like to address the second question proposed above, to determine whether items present in the buffer exist in a privileged state such that they will be more easily retrievable than items not present in the buffer during time of test, and will not be influenced by other items on the list.

2.6 List Length Manipulations

Experiment 2 indicates that recall patterns seen in immediate free recall are due to retrieval of items from a privileged buffer state. If recency is due to retrieval from the rehearsal buffer, similar recency effects should be apparent regardless of list length for all lists that exceed the size of the buffer (Murdock, 1962). Patterns of recency, as evident in SP and FRP effects, should be similar for single item lists of all lengths > n and similar for paired item lists of all lengths > n, where n refers to the size of the buffer.

Additionally, while the magnitude of CRP effects may change with list length, we would expect to see consistency, such that more +1 lag transitions occur for paired items than single items, regardless of list length.

The critical test of the model relates to first recall probabilities. In the previously described experiment, we observed differential first recall probabilities for single versus paired item lists. For single item lists, the last item on the list is the most likely item to be recalled first, whereas for paired item lists, the second to last item on the list (or first item of the last pair) is most likely to be recalled first. If these patterns are due to retrieval from a rehearsal buffer, then additional items studied should not influence the items that are currently present in the buffer at time of test, and the same first recall patterns should be seen regardless of list length. If, however, the memory system does not include a rehearsal buffer, then first recall probabilities should be influenced by list length because the most recently studied items would suffer different amounts of interference in lists of different lengths.

The model is constrained by data from all of the previous work discussed in this manuscript. Thus, we present predictions related to list length manipulations using the same set of parameters used in the previously described models. It is necessary, however, to specify the parameters of the model that will be affected by list length. First, it is assumed that more attempts at sampling and recovery of items will be made for longer list lengths than shorter list lengths, thus the stopping rule is a function of list length, where the number of sampling and recovery attempts = $(List\ Length * \kappa)$. Next, we assume that as list length increases (and context changes), it will be harder to reinstate context features from the beginning of the list, and as it becomes harder to reinstate these features, the probability of sampling from the buffer becomes more likely. Thus, the probability of reinstating beginning of list context features is also a function of list length:

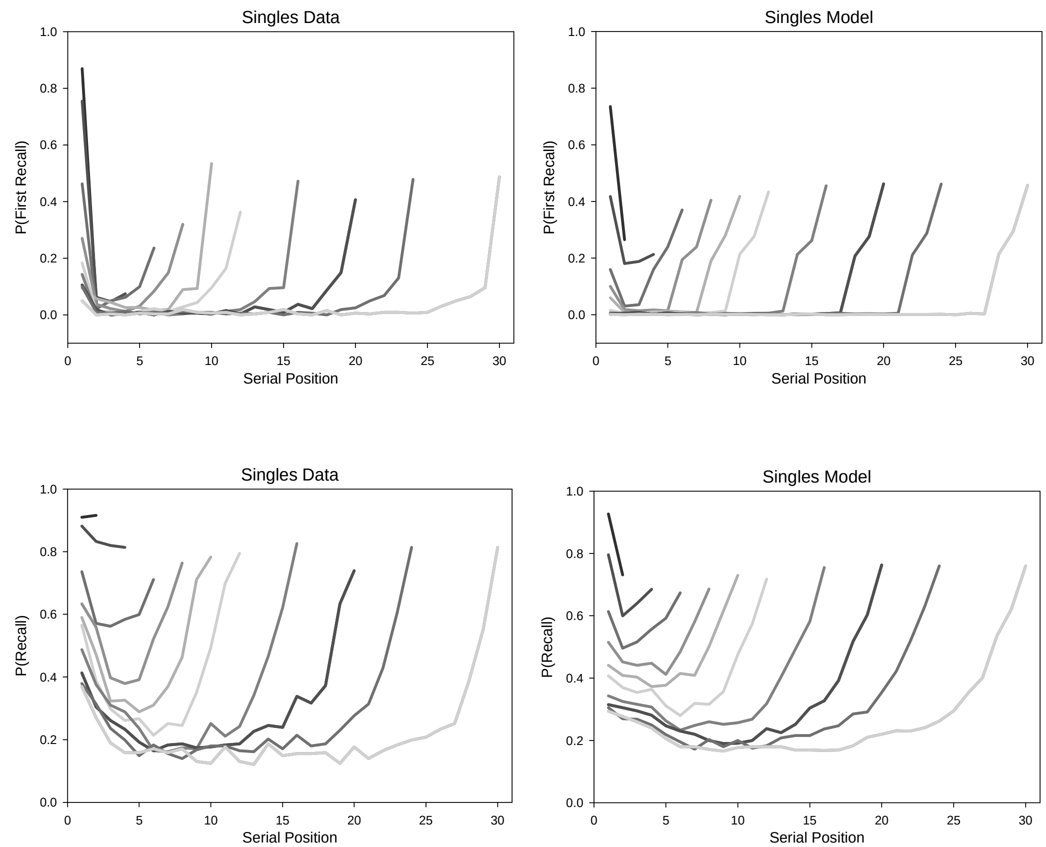
$\gamma = (ListLength)^{-1}$,

and the probability of initially sampling from the buffer is equal to 1 - γ. The model predictions are shown in the right panels of Figures 4, 5, and 6 for FRPs, SPs, and CRPs, respectively.

As with Experiment 2, the model predicts differential patterns of FRPs and CRPs for single versus paired items. Importantly, it predicts that patterns of FRPs should not change with list length (for all list lengths greater than four). Additionally, it makes clear predictions about zigzag effects in the SP curves. For paired items, we see a zigzag pattern in both SP and FRP curves. While the penultimate item is most likely to be recalled first, items from earlier in the list are sometimes recalled first; however the first item in a pair is always more likely to be recalled first than the second item in a pair. These effects are consistent for all list lengths (greater than four).

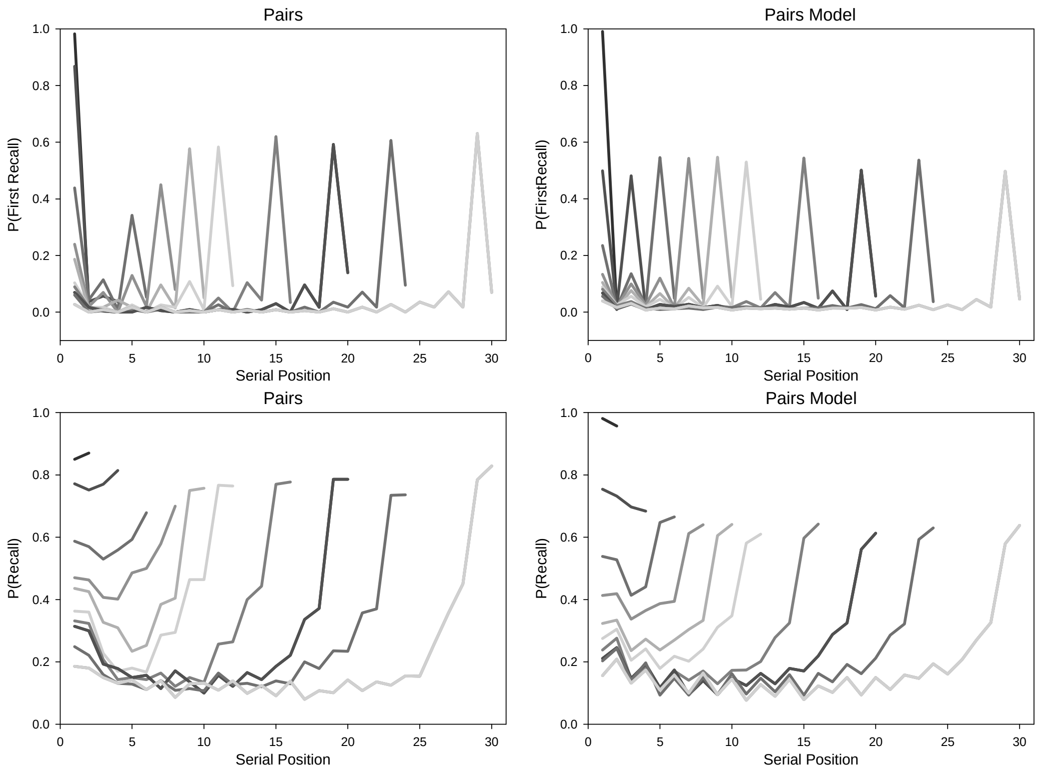
A second test of the model involves reaction times in free recall. As suggested by Davelaar et al. (2005), if retrieval in immediate free recall begins by sampling items from the buffer, then response time to output the first item during recall should not be influenced by other items on the list that are not present in the buffer. Thus, a model that includes retrieval from the buffer predicts that time delay to output the first item during recall should not be affected by list length, whereas a model that does not include retrieval from the buffer would predict that a memory search should include all items rather than just items that exist in a privileged state, thus response times should be longer

Figure 12: Single item list length data and model predictions for first recall probabilities and serial position effects



to output the first item from a long list compared to a short list. While the Lehman-Malmberg model is not a model of reaction time, these predictions are consistent with those of other models of reaction time (Davelaar et al., 2005); such assumptions that have yet to be built into the current model in order to account for reaction time data. However, this prediction is based on the assumption that reaction times are a function of the probability of sampling a given item compared to the probability of sampling all items in a retrieval set; when there are fewer items in the retrieval set, as is the case with sampling from a buffer, the probability of sampling a given item will be greater than when there are more items in a retrieval set. Thus, if the first attempt at retrieval is restricted to the

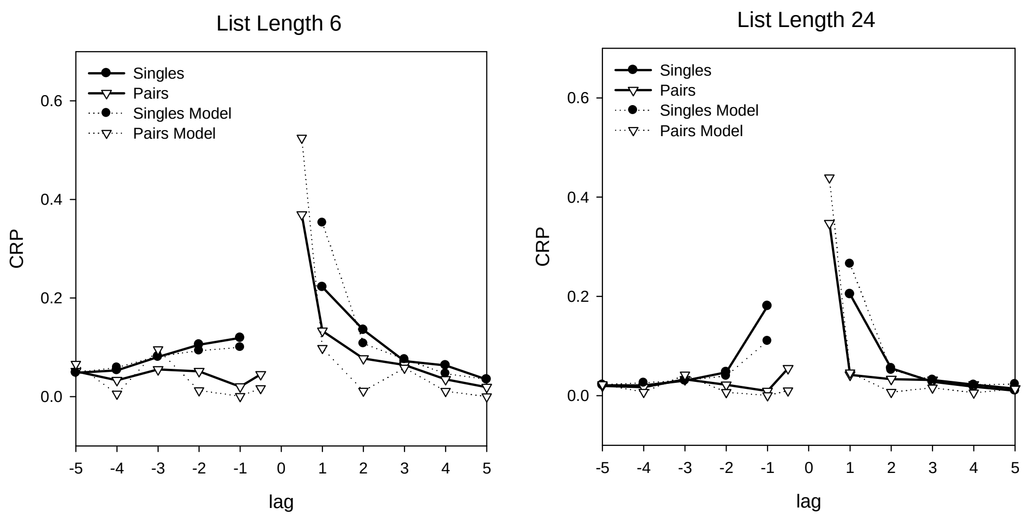
Figure 13: Paired item list length data and model predictions for first recall probabilities and serial position effects



items in a buffer, then the probability of sampling a given item in the buffer would not be affected by list length, thus reaction time should be consistent across list lengths. If, however, retrieval does not occur from a buffer, then the likelihood of sampling a given item would be decreased, predicting a greater reaction time for items from longer lists.

To address the issue of the necessity of a buffer in the retrieval process, we conducted a second experiment utilizing single-item and paired-item study lists, where list length was manipulated (cf. Ward, Tan & Grenfell-Essam, 2010), and examined SPs, FRPs, and CRPs, in addition to reaction time to output the first item recalled for each list. The examination of these effects allows us to test the model's predictions regarding retrieval from a privileged buffer state.

Figure 14: Conditional response probability data and model prediction for single and paired item study lists of length 6 and 24



2.7 Experiment 3

2.7.1 Method

2.7.1.1 Participants, Materials, and Procedure. Participants were 176 undergraduate psychology students at the University of South Florida who participated in exchange for course credit (as all variables were manipulated within subject, it was necessary to collect data from many participants in order to eliminate noise and see clear serial position effects). For each participant, two word lists of each list length (2, 4, 6, 8, 10, 12, 16, 20, 24, and 30 items) were created for each study condition (single-item lists and paired-item lists), each consisting of randomly related concrete nouns (between 20 and 50 occurrence per million; Francis and Kucera, 1982). Thus, for each participant, a total of 40 word lists were created, half for the single-item study list condition and half for the paired-item study condition. Item presentation occurred in the same manner described in Experiment 2 for single and paired lists, with similar study times. In order to reduce the effects of fatigue, the experiment was run in two sessions, a week apart, so that participants completed only 20 study-test cycles in each session. In one session, participants completed all study-test cycles for paired-item lists, and in the other they completed all study-test cycles for single-item lists (with the order of single and paired conditions counterbalanced). In each study-test cycle, the lists were randomly presented, such that the length of each new list was not predictable, and participants were not told in advance the length of each list.

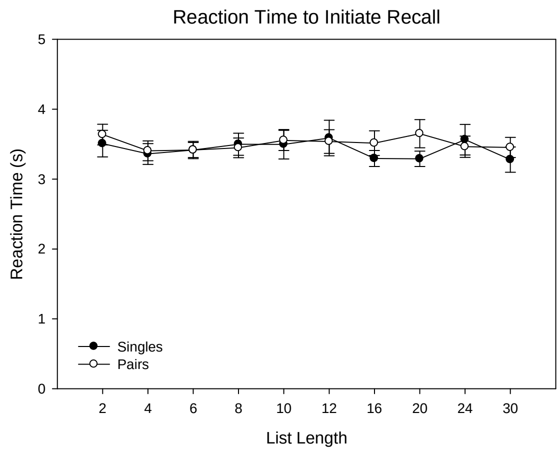
2.7.2 Results

As the focus of this experiment is the qualitative effects of buffer operations and their result on serial position curves as list length increases, we primarily focus on qualitative patterns visible in the data rather than quantitative statistical comparisons (the examination of a serial position by list length interaction is not possible, as each list length has a different number of serial position points). As shown on the left panels of Figure 12, there is a shift from primacy toward recency as list length increases for both first recall probabilities and serial position effects; for longer lists, participants are more likely to begin recall with the last item on the list, and more likely to recall items from the end of the list than the beginning of the list. The left panels of Figure 13 reveal that, as in Experiment 2, the serial position curve and first recall probabilities look quite different for paired item study lists. Whereas participants are likely to begin recall with the last item on the list for long lists of single items, participants are most likely to begin recall with the penultimate item on the list for paired items $t(158) = 7.64$, $SE = .04$, $p < .001$. As seen in the top left panel of Figure 13, this pattern is consistent for all list lengths greater than four. Additionally, for all list lengths we see a zigzag pattern in first recall probabilities; items that are the first member of a pair are more likely to be first recalled than items that are the second member of a pair (for which the probability of first recall is almost zero). This was confirmed by a Chi-square test comparing the likelihood of first recalling the first member of a pair versus the second member of a pair for a list of 24 items, excluding the first and last pairs on the list, $\chi^2(1) = 16.71 > 3.84$. Further, the zigzag pattern is consistent throughout recall, and this does not occur for single items, $\chi^2(1) = 3.00 < 3.84$. The bottom panel of Figure 13 shows that throughout the list, when a first item from a pair is recalled, the second item from that pair is recalled with almost equal probability, and again this is not the case for single items.

Figure 14 shows conditional response probabilities for two list lengths, one short (6 items) and one long (24 items). For both the short list and the long list, we see the typical patterns in the CRPs – a greater likelihood to transition to a nearby serial position, and an asymmetry in that recall is more likely to move in the forward direction. For both list lengths, there is a greater likelihood of moving forward one item within a pair than moving forward one item for single items (all $p < .05$).

Finally, we examined reaction time to output the first item recalled. This was measured from the time the test began until the participant pushed “Enter” to submit the word. Participants whose mean response times were greater than three standard deviations from the mean were removed from these analyses. As shown in Figure 15, reaction time to output the first item recalled was consistent across all list lengths. Reaction times did not differ between single and paired item lists, and there was no effect of list length on reaction time or no interaction of condition and list length, all $p > .05$.

Figure 15: Reaction time data for single and paired item study lists of each length



2.7.3 Discussion

These data give support to the theory that in immediate free recall, initial retrieval occurs from a privileged buffer state. Given the zigzag pattern in first recall probabilities in pairs, it is apparent that recall does not always begin with the most recent items on the list; however the greater likelihood of first recalling the last item on the list for single items and the penultimate item on the list for paired items is consistent for all list lengths (greater than four) suggests that once the list length exceeds the buffer size, recent items will be maintained in a privileged state. Finally, reaction time to output the first item during recall did not increase with list length, suggesting that the items that are output first are not suffering interference from other items on the list.

In general, the model does a good job of fitting the data. It predicts the correct patterns of SPs and FRPs across list lengths. The model also produces good qualitative fits for the CRPs, where lag +1 transitions are more likely for pairs then for single items. Quantitatively, the model is overpredicting such transitions, especially for pairs. This is mainly due to the simplifying assumptions that we made related to chunking processes – context is only stored for the first item in a pair, and if the first item in a pair is retrieved, the next item sampled will always be the second item from a pair. Introducing some variation in both of these components, such that context is weakly stored for the second item in a pair, and other items are allowed to be sampled after the first item in a pair is retrieved would likely lead to more accurate quantitative fits. However, at this time, we are more interested whether the model is able to produce the qualitative patterns of data in these conditions, which it does quite well.

It should be noted that while recall patterns were similar for all list lengths greater than four, these patterns differed for short lists, i.e. lists less than four items long. This finding is consistent with prior work showing that for short lists, subjects typically begin recall with items at the beginning of the list rather than the end of the list as they do with long lists (Ward et al., 2010). As our intention is to examine predictions regarding the privileged state of items present in the buffer, we are more interested in model predictions at longer list lengths. The model fits the data less well for lists of lengths two and four, because it includes no assumptions about different retrieval processes used for short lists. Additional assumptions would be necessary in order to fit the data for short lists. For example, one might assume that when the capacity of the buffer has not yet been reached, participants rely solely on reinstating the beginning of list context as a cue rather than relying somewhat on current context.

Chapter 3

General Discussion

The challenge for this model was to account for all of the data from the previously discussed paradigms, including intentional forgetting, free recall and recognition, immediate and delayed free recall, the continuous distractor task, and Craik & Watkins (1973) data, in addition to the data from the experiments in the single/pairs paradigm with differing list lengths, using the same set of parameter values. In this sense, the model is able to handle a large number of data points with very few free parameters.

In order to determine whether our model provides the best possible fit to the data from the various experiments described in this manuscript, we compared the predictions of the version of the model described in this manuscript to those generated from similar versions of the model which differed only in the use of the buffer.

3.1 Buffer as an Encoding Process

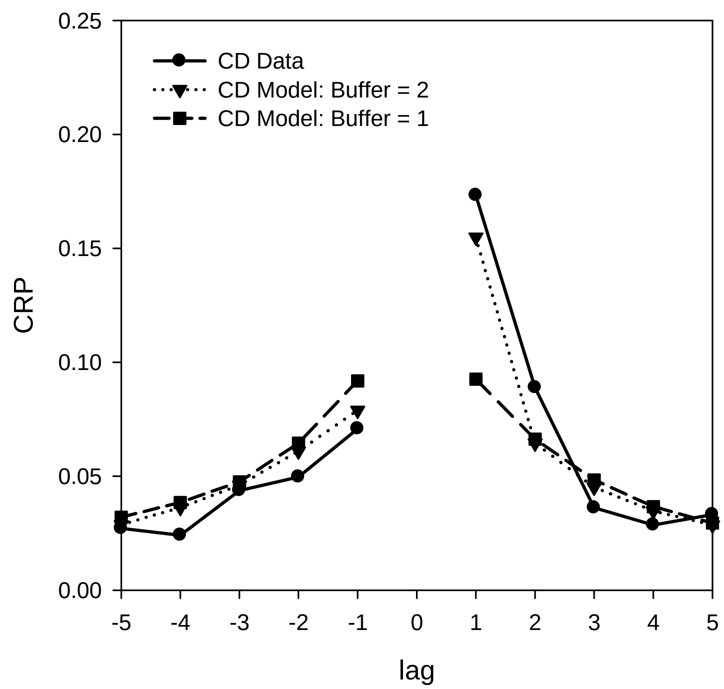
The first question we were interested in addressing was whether the buffer provides a good model of encoding processes. We showed that our model was able to fit the data from a continuous distractor task, which has been said to be troublesome for dual-store models. Our model accounts for this data by assuming that the buffer size is reduced from two to three items as a result of the continuous distractor task, however we assume that two items can be simultaneously maintained despite this task. In order to show that the use of the buffer is necessary in order to obtain the model fits presented above, we compared the model to one in which the buffer was not used at all (i.e. a single-store model). The results of model simulations using this model are presented in Figure 16. While the model is able to provide qualitatively accurate fits of the data for FRPs and SPs, it fails in predicting the forward asymmetry seen in the CRPs, as it relies on context retrieved from a recalled item as the retrieval cue, which has an equal likelihood of sampling a previous item or a subsequent item.

Thus, this model is at least equally able to account for long-term recency effects and contiguity effects as single-store models, without the use of any additional model parameters. While this alone does not suggest that single-store models are wrong, it does tell us that dual-store models can do the job, in addition to accounting for a variety of other data. Additionally, our model is able to fit new data that will be challenging for single-store models that don't incorporate control processes to explain.

3.2 Buffer as a Retrieval Process

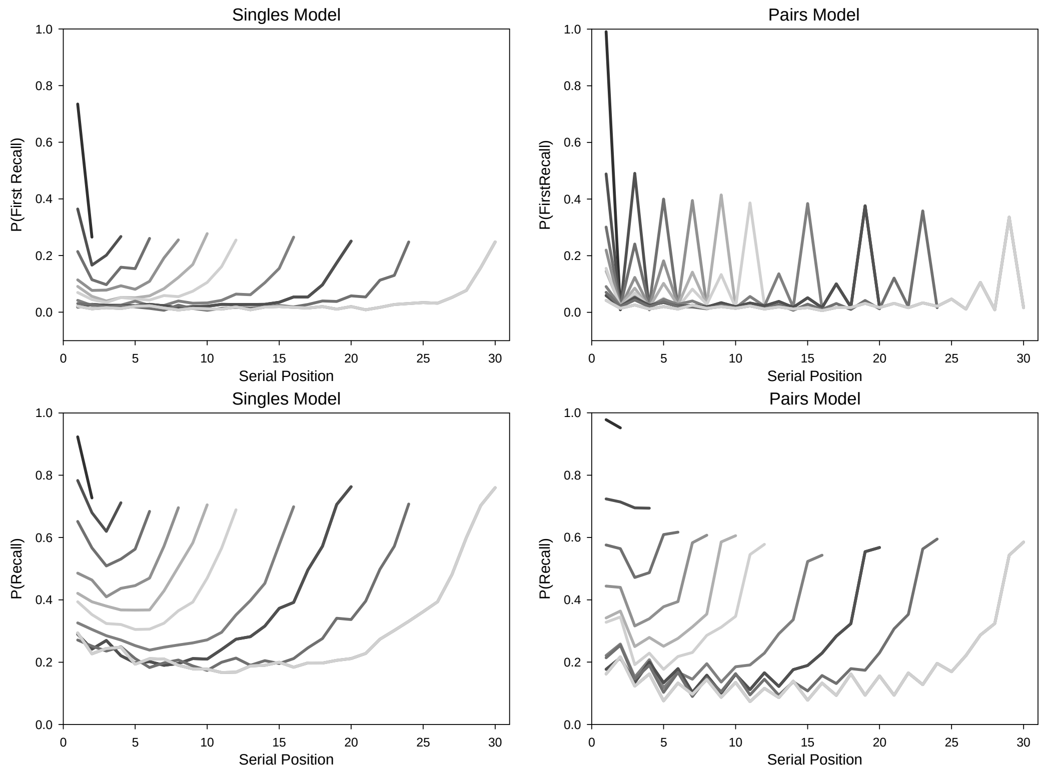
The second question we addressed concerns the use of the buffer as a retrieval process. In order to assess the necessity of such a process, we compared our model to a version in which associations are made between items in the buffer, but these items do not exist in a privileged state during retrieval, such that items are matched to the current context cue, but any item from the list may be initially sampled (rather than any of the items presently in the buffer at this time). These model simulations are shown in Figure 17. In sum, both models provide good qualitative fits to the data in that they both predict the penultimate item peaks in FRPs for paired lists, in addition to the zigzag patterns

Figure 16: Model predictions for conditional response probabilities using different buffer sizes in continuous distractor free recall



throughout the list in SPs and FRPs for paired lists. While both models provide good qualitative accounts of the data, the goodness of fit test reveals that neither provides a good quantitative fit when measured by the chi square test. We used a chi square criterion value of 124.3, which corresponds to 100 degrees of freedom because most tables don't include chi square values for more than 100 degrees of freedom; however our true degrees of freedom (528) far exceed this due to the number of parameter values we are fitting. Thus, this test may not be appropriate for determining acceptable fits in a model that includes so many data points. It can, however, show us which model provides a better fit to the data. In this case, the model that includes retrieval from the buffer, however provides a better quantitative fit to the data than the model that does not include

Figure 17: Model without retrieval from the buffer



retrieval from the buffer, X^2 (528) = 2989 and 4248, respectively. While it may be possible to improve the fit of the non-buffer retrieval model, this would require additional parameters, increasing the complexity of this model. Additionally, further evidence for retrieval from the buffer comes from the reaction time results in the list-length study. For both single and paired-item study lists, reaction times are consistent across list lengths, suggesting that the items that are first output are present in a state that is not subject to interference from other items on the list. A further model test will involve examining the model's predictions for reaction time rather than recall processes. Such a model is beyond the scope of the current manuscript, but will be the subject of future work.

3.3 Comparing the Lehman-Malmberg Model to Other Models

3.3.1 TCM

According to TCM (Howard & Kahana, 2002), context drifts gradually over time. While in the Lehman-Malmberg model context drifts in a random manner (Mensinck & Raaijmakers, 1989), context drifts nonrandomly in TCM. When items are being studied, contextual drift is driven by the retrieval of preexperimental contextual states associated with those items. During recall, both the preexperimental and studied contexts are retrieved with a recall item, which drives the evolution of context during test (whereas the Lehman-Malmberg model utilizes retrieved study context and item information). Thus, TCM includes the use of a single memory store, and does not include control processes utilized by the subject either during encoding or retrieval. Howard and Kahana have argued that SP, FRP, and CRP functions are not the result of a rehearsal buffer; rather they are the result of the use of the current state of context to probe memory, combined with the contextual drift process that occurs during encoding and retrieval. While they do not deny that control processes may play a function in memory, they argue that recency and contiguity effects are the result of a basic memory process (i.e. out of an individual's control), and not the result of buffer operations. We argue, on the other hand, that buffer operations are necessary in order to account for the wide variety of data that can be handled by the Lehman-Malmberg model. Howard and Kahana also suggest that while contextual encoding processes in their model may mimic the functions of the buffer, items do not exist in the buffer in an all-or-none fashion. The data from the experiments reported in this manuscript indicate, however, that items at the end of a list are present in a privileged buffer state and do not suffer interference from items not present in the buffer.

3.3.2 Activation Models

Davelaar and colleagues (Davelaar et al., 2005; 2006) have developed a dual-store neurocomputational model of memory that includes both short-term and long-term memory components with a foundation in neuropsychological processes. This model proposes a short-term buffer and a long-term memory store both as structural components, linked to specific neuroanatomical processes of excitation and inhibition in the brain. Whereas the Lehman-Malmberg model and TCM are models in which retrieval of an item occurs based on the match between a retrieval cue and the item, the activation model specified by Davelaar et al. conceives of retrieval as a result of item activation, where items currently in the buffer are all activated, and items in long-term memory may or may not be activated at a given time. Activation in long-term memory is based on episodic context matching, but, in contrast to the Lehman-Malmberg model, retrieval from the short-term buffer is based on current activation levels, and items in the buffer are deactivated via inhibition from other items that have entered the buffer, rather than simply displaced by the new items. Thus, rather than a flexible control system that can be differentially used depending on the way information is rehearsed, the buffer in their model is simply a set of the most recent traces which have activation levels above a threshold; such activation levels are a function of the structural memory system and are not under the control of an individual.

While our data do not provoke us to make a strong argument *against* activation models, we believe our model provides more consistency with past work. For example, it is possible to maintain a small number of items in a rehearsal buffer while processing unrelated sentences (Baddeley & Hitch, 1974). Such findings are consistent with the idea that the rehearsal buffer is a flexible control process that can be adapted to a task, and challenging to models that conceive of buffer as a structural short-term memory store. While Davelaar et al. (2005) don't include control processes in their model, they do address the ways in which such processes may contribute to the model. However, at this time, their model has not been used to account for rehearsal processes during encoding and serial order information that results from the use of the buffer as a control process during encoding.

3.4 Utilizing the Buffer in Chunking Operations

A defining characteristic of the Lehman-Malmberg model is the way in which the buffer may be used to allocate encoding resources in different ways, depending on the nature of the task, a feature borrowed from the A-S model (Atkinson & Shiffrin, 1968). For example, if a larger buffer size is being used, and a single item is present in the buffer, this allows for more contextual encoding for that item than if a small buffer size is used. As more items enter the buffer, encoding resources are distributed so that associations between items are also stored. Some tasks may allow an individual to use a larger buffer size, or they may influence the way items are dropped from the buffer. Other tasks will affect whether or not retrieval from the buffer is possible.

The flexibility of the rehearsal buffer allows it to account for chunking operations in memory. For example, when studying pairs of items, subjects may use the buffer to associate a large amount of contextual information with the first item in a pair and a minimal amount of contextual information with the second item in a pair. Further, strong item to item associations may be made between items within a chunk so that this information can be use to produce entire chunks of information during recall. It is not clear how a model without such control processes would be able to account for the effects of chunking on recall.

While other memory models such as TCM (Howard & Kahana, 2002) and Davelaar et al.'s (2005) activation model do not argue against control processes, these models do not include any such processes, which we argue are critical in accounting for the data described in this manuscript. These control processes are necessary in order to produce the patterns of retrieval shown in the SP, FRP, and CRP curves reported in the current experiments. The utilization of the buffer as a control process also allows the model to account for data that has been said to be troublesome for buffer models in the past.

3.5 Extending the Model

One potential limitation of this work is that we typically use lists of words as study materials, thus it is not clear how this model may generalize outside of verbal stimuli. This model has been extended to account for memory of both words and pseudowords (pronounceable nonwords; Lehman & Malmberg, unpublished data), and similar models have also been used to fit data from experiments using faces, pictures, or other visual stimuli, such as Chinese characters (Annis & Malmberg, unpublished data; Xu & Malmberg, 2007). Thus, we have confidence that our models generalize to various types of stimuli, including both verbal and visual stimuli, and semantic and nonsemantic stimuli. There are many advantages of using wordlists; they can be highly controlled, easy to present, and easy to test for in various types of memory tasks. It may be useful to examine our model in the context of a more naturalistic paradigm; however this is beyond the scope of the present project. Regardless, the stimuli used in experimental memory tasks are seen as an analogue to many of those encountered in everyday life, and we assume that our findings will generalize. For example, something that someone might try to remember in everyday life might be a grocery list or which items to pack for a trip, and we assume that one would make use of similar mechanisms to accomplish such tasks.

The goal in developing this model is to create a complete account of various memory processes. The model is able to account for performance on various types of memory tasks, including free recall and recognition, in addition to performance that results from many different encoding strategies, ranging from incidental learning to intentional learning to intentional forgetting. Thus, we now have a more comprehensive account of encoding and retrieval processes, addressing some of the issues that have challenged dual-store models of memory for the past 40 years.

3.5.1 Individual Differences

This model is used to generate average predictions, for data that is collapsed over many subjects. One of the short-term goals for this model is to begin to account for individual differences in cognitive processes. Typically in cognitive research, we look for group differences in the effects of our independent variables, but rarely use correlational analyses to inform us about why some individuals experience certain cognitive patterns and others do not (Widaman, 2008; see also McDowd & Hoffman, 2008). This model may be extended to account for patterns of cognition seen in special populations and can allow us to evaluate theories about what leads to these patterns. Further, the model can be used to generate testable predictions related to such patterns. For example, one area of research to which this model may be applied is in the study of cognitive deficits in depression.

There is a large body of research related to cognitive deficits in depression, much of which suggests that depressed individuals show impairments in various cognitive processes, including concentration, attention, and memory (American Psychiatric Association [*DSM-IV-TR*], 2000; Burt, Zembar & Neiderehe, 1995; Christopher & MacDonald, 2005; Cohen, Weingartner, Smallberg, Pickard & Murphy, 1982; Kalska, Punamaki, Pelli, & Saarinen, 1999), or that they have negative information processing biases (Blaney, 1986; Bradley & Mathews, 1983; Matt, Vasquez, & Campbell, 1992).

Recent interest has developed in intentional forgetting processes in depression, driven by the suggestion that intentional forgetting is related to inhibitory processes, and that inhibitory processes are impaired in individuals with depression (Johnson, 2007; Joormann, Yoon & Zetche, 2007; Power, Dalgleish, Claudio, Tata, & Kentish, 2000). If this is true, then depressed individuals would be expected to show deficits on intentional forgetting tasks, particularly when negative materials, toward which depressed individuals may be negatively biased to attend are involved.

Lehman and Malmberg (2009) proposed the context differentiation model for intentional forgetting; the model provides a formal account of the process commonly referred to as inhibition which causes the effects of intentional forgetting. Lehman and Malmberg (2011) proposed that impairments seen on directed forgetting tasks in depressed populations may be the result of a failure to use context differentiation in order to compartmentalize information that is not relevant to the task at hand, a problem that may also manifest in the symptoms of depression, such as ruminating on depressive thoughts when these thoughts are not presently useful.

The Lehman-Malmberg model can be utilized to make formal predictions about what will occur in various tasks for individuals who are not able to use context differentiation in order to complete the compartmentalization process. These predictions can then be tested in order to evaluate the viability of this theory. For example, we can generate predictions based on simulations using trials where contextual differentiation is used and compare these to those for trials where contextual differentiation is not possible, and trials where contextual differentiation is forced (for example, through the use of changes in environmental context). We can then evaluate these predictions by generating data from depressed participants and comparing these data to our model's prediction (a task that is currently underway).

3.5.2 Neuroscience Models

One future direction for this model is to take on these issues from a neural perspective (Cowan, 1995; Widaman, 2008). Dissociations in neuropsychology are often cited as evidence of separate short-term and long-term memory stores (Davelaar et al., 2005). However, Crowder (1989) suggests that one type of amnesia may be due to a specific type of coding deficit in which the relations of items to their temporal contexts are not properly coded. A model that is able to account for biological measures in addition to behavioral measures would be useful in evaluating the dual-store issue. Additionally, it would allow us to make a greater variety of predictions with the model. For example, depressed individuals show reduced activity in the prefrontal cortex

(Henriques & Davidson, 1991), an area that is associated with various cognitive processes (Posner, 1992). Developing a comprehensive model of both brain and behavioral processes would allow us to further explore the applications of this model in special populations. One advantage of the Davelaar et al. (2005) model is that it incorporates features of the neural processes of excitation and inhibition in order to generate activation levels in memory. Future work will attempt to accomplish similar goals in the framework of the Lehman-Malmberg model. Perhaps the assumptions of these models can be combined in a way that provides a model with the advantages of the Lehman-Malmberg model and the Davelaar et al. model in order to account for a much larger variety of data than can currently be handled by either model alone.

3.6 Conclusion

In sum, the Lehman-Malmberg model successfully accounts for a variety of data in multiple episodic memory paradigms. It has been shown to make accurate predictions related to both intentional and unintentional forgetting, maintenance rehearsal, continuous distractor tasks, and chunking operations. The key characteristics that allow the model to fit this wide array of data are the operations of the buffer as a process during both encoding and retrieval. At present, our findings suggest that control processes, which we conceptualize as a rehearsal buffer in the Atkinson and Shiffrin (1968) tradition, are a necessary component of episodic memory.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (Revised 4th ed.). Washington, DC: Author.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 261-295.

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach & T. Carr (Eds.), *Inhibitory processes in attention, memory, and language (pp. 265-325)*. New York: Academic Press.

Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In Spence, K.W. & Spence, J.T. (Eds.), *The Psychology of Learning and Motivation (pp 89-195),* New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971). The Control of Short-Term Memory. *Scientific American, 225*(2), 82-90.

Baddeley, A.D., Hitch, G.J.L (1974). Working Memory, In G.A. Bower (Ed.), *The psychology of learning and motivation: advances in research and theory (*Vol. 8, pp. 47–89), New York: Academic Press.

Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 603–616.

Basden, B. H., Basden, D. R., & Wright, M. (2003). Part-list re-exposure and release of retrieval inhibition. in directed forgetting. *Consciousness and Cognition, 12*, 354-375.

Benjamin, A. S. (2006). The effects of list-method directed forgetting on recognition memory. *Psychonomic Bulletin & Review.*

Bjork, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory (pp. 217-235).* Washington, D.C.: Winston.

Bjork, R. A. (1978). The updating of human memory. In G. H. Bower (Ed.), *The psychology of learning and motivation. (Vol. 12., pp. 235-259*). New York: Academic Press.

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition, 5,* 176-196.

Bjork, R. A., & Geiselman, R. E. (1978). Constituent processes in the differentiation of items in memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 347-361.

Bjork, R.A., LaBerge, D., & Legrand, R. (1968). The modification of short-term memory through instructions to forget. *Psychonomic Science, 10,* 55-56.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6*(2), 173–189.

Blaney, P.H. (1986). Affect and memory: A review. Psychological Bulletin*, 99*(2), 229-246.

Block, R. A. (1971). Effects of instructions to forget in short-term memory. *Journal of Experimental Psychology, 89 ,* 1–9.

Bradley, B. & Mathews, A. (1983). Negative self-schemata in clinical depression. *British Journal of Clinical Psychology, 22*(3), 173-181.

Broadbent, D. E. (1958). *Perception and comunication*. New York: Academic Press.

Burt, D. B., Zembar, M. J., & Niederehe, G. (1995). Depression and memory impairment: A meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin, 117,* 285–305.

Christopher, G. & MacDonald, J. (2005). The impact of clinical depression on working memory. *Cognitive Neuropsychiatry*, 10(5), 375-379.

Cohen, R. M., Weingartner, H., & Smallberg, S. A., Pickard, D., & Murphy, D. L. (1982). Effort and cognition in depression. *Archives of General Psychiatry, 39*, 593-597.

Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-185.

Cowan, N., Wood, N. L., & Borne, D. N. (1994). Reconfirmation of the short-term storage concept. *Psychological Science, 5*(2), 103-106.

Criss, A.H. and Shiffrin, R.M. (2004). Context noise and item noise jointly determine recognition memory: A comment on Dennis & Humphreys (2001). *Psychological Review,* 111(3), 800-807.

Craik, F. I. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning & Verbal Behavior, 9*(2), 143-148.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11*(6), 671-684.

Craik, F. I., & Watkins, M. J. (1973). The role of rehearsal in short-term memory1. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 599–607.

Crowder, R.G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Crowder, R. G. (1989). Modularity and dissociations in memory systems. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving*. (pp. 271-294). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition*, *21*(2), 142-145.

Dalezman, J. J. (1976). Effects of output order on immediate, delayed, and final recall performance. *Journal of Experimental Psychology: Human Learning and Memory, 2*(5), 597-608.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review, 112*(1), 3–41.

Davelaar, E. J., Haarmann, H. J., Goshen-Gottstein, Y., & Usher, M. (2006). Semantic similarity dissociates short-from long-term recency effects: Testing a neurocomputational model of list memory. *Memory & Cognition*, *34*(2), 323.

Deese, J., & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology, 54*(3), 180–187.

Dennis, S. & Humphreys, M. S. (2001). A Context Noise Model of Episodic Recognition Memory. *Psychological Review*. 108, 452-478.

Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology*, 16, 519-555.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. New York: Teachers College, Columbia University.

Eich, J. E., Weingartner, H., Stillman, R. C., & Gillin, J C. (1975). State-dependent accessibility of retrieval cues in the retention of a categorized list. *Journal of Verbal Learning & Verbal Behavior, 14*(4), 408-417.

Elmes, F. J., Adams, C., & Roediger, H. L. (1970). Cued forgetting in short-term memory: Response selection. *Journal of Experimental Psychology, 86,* 103–107.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, 62, 145–154.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Geiselman, R. E. & Bagheri, B. (1985). Repetition effects in directed forgetting: Evidence for retrieval inhibition. *Memory and Cognition, 13,* 57-62.

Geiselman, R. E., Bjork, R. A., & Fishman, D. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112,* 58–72.

Gillund, G. & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1-67.

Glanzer, M. & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning & Verbal Behavior, 5*(4), 351-360.

Glenberg, A. M. (1984). A retrieval account of the long-term modality effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(1), 16–31.

Gobet, F., Lane, P. C. R., Croker, S., C-H Cheng, P., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*(6), 236-243.

Godden, D., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology, 66,* 325–331.

Goldstein, E.B. (2008). *Cognitive Psychology: Connecting Mind, Research, and Everyday Experience*, 2nd edition. Belmont, CA: Wadsworth

Goodwin, D. W., Powell, B., Bremer, D., Hoine, H., & Stern, J. (1969). Alcohol and recall: State dependent effects in man. *Science, 163,* 1358–1360.

Henriques, J.B. & Davidson, R. J. (2000). Deceased responsiveness to reward in depression. *Cognition and Emotion*, 14(5), 711-724.

Howard, M. W., & Kahana, M. J. (1999). Contextual Variability and Serial Position Effects in Free Recall. Journal of Experimental Psychology: *Learning, Memory, and Cognition, 25*(4), 923–941.

Howard, M.W. & Kahana, M.J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*, 269-299.

Humphreys, M. S. (1976). Relational information and the context effect in recognition memory. *Memory & Cognition*, 4, 221-232.

Humphreys, M.S., Bain, J.D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks. *Psychological Review*, 96, 208-233.

Jacoby, L. L (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513-541.

James, W. (1890). *The principles of psychology*. New York: Henry Holt and Company.

Jang, Y. & Huber, D.E. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34*(1) 12-27.

Johnson, N. F. (1970). The role of chunking and organization in the process of recall. *The psychology of learning and motivation: Advances in research and theory, 4,* 171–247.

Johnson, S. L. (2007). Cognitive inhibition across psychopathologies. *Applied and Preventive Psychology, 12*(3), 97–98.

Joormann, J., Yoon, K. L., & Zetsche, U. (2007). Cognitive inhibition in depression. *Applied and Preventive Psychology, 12*(3), 128–139.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, 24, 103–109.

Kalska, H., Punamäki, R., & Mäkinen-Pelli, T. (1999) Memory and metamemory functioning among depressed patients. *Applied Neuropsychology, 6*(2), 96-107.

Koppenaal, L., & Glanzer, M. (1990). An examination of the continuous distractor task and the 'long-term recency effect.'. *Memory & Cognition, 18*(2), 183-195.

Lehman, M., & Malmberg, K. J. (2009). A global theory of remembering and forgetting from multiple lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 970.

Lehman, M., & Malmberg, K. J. (2011). Overcoming the effects of intentional forgetting. *Memory & Cognition*, *39*(2), 335-347.

Light, L. L., & Carter-Sobell, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and Verbal Behavior, 9,* 1-11.

Loft, S., Humphreys, M.S., & Whitney, S.J. (2008). Control of access to memory: The use of task interference as a behavioral probe. *Journal of Memory and Language, 58*(2), 465-479.

Macht, M. L., Spear, N. E., & Levis, D. J. (1977). State-dependent retention in humans induced by alterations in affective state. *Bulletin of the Psychonomic Society, 10,* 415–418.

MacLeod, C.M. (1975). Long-term recognition and recall following directed forgetting. *Journal of Experimental Psychology: Human Learning and Memory*, *1*, 271–279.

MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1–57). Mahwah, NJ: Erlbaum.

Macleod, C. M, Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition to inhibition. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (pp. 163-214). New York, NY: Elsevier Science.

Malmberg, K. J. (2008). Towards an understanding of individual differences in episodic memory: Modeling the dynamics of recognition memory. in A. Benjamin and B.

Ross (Eds.) *The Psychology of Learning and Motivation: Skill and Strategy in Memory Use*, Vol. 48, 313-349.

Malmberg, K. J., Holden, J. E., Shiffrin, R. M. (2004). Modeling the Effects of Repetitions, Similarity, and Normative Word Frequency on Old-New Recognition and Judgments of Frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 319-331.

Malmberg, K. J. & Murnane, K. (2002). List composition and the word-frequency effect for recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 616-630.

Malmberg, K. J., Lehman, M., & Sahakyan, L. (2006). On the Cost and Benefit of Taking it out of Context: Modeling the Inhibition Associated with Directed Forgetting, *Proceedings of the 28th Meeting of the Cognitive Science Society,* 549-554.

Malmberg, K. J. & Shiffrin, R. M. (2005). The 'One-Shot' Hypothesis for Context Storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 322-336.

Malmberg, K. J. & Xu, J. (2007). On the flexibility and on the fallibility of associative memory, *Memory & Cognition*, 35(3), 545-556.

Malmberg, K. J., Zeelenberg, R., & Shiffrin, R.M. (2004). Turning up the Noise Or Turning Down the Volume? On the Nature of the Impairment of Episodic Recognition Memory by Midazolam. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 540-549.

Matt, G. E., Vazquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review, 12,* 227–255.

McDowd, J. M., & Hoffman, L. (2008). Challenges in attention: Measures, methods, and applications. In S. M. Hofer & D. F. Alwin (Eds.) *Handbook of cognitive aging: Interdisciplinary perspectives*, (pp. 122-133). Los Angeles: Sage Publications.

Mensink, G. J. M., & Raaijmakers, J. G. W. (1989). A model for contextual fluctuation. *Journal of Mathematical Psychology*, 33, 172–186.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97.

Murdock, B. B. (1962). The serial position effect of free recall*. Journal of Experimental* Psychology, 64(5), 482–488.

Murdock, B. B. (1995). Primacy and recency in the chunking model. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension: Essays in honor of Walter Kintsch.* (pp. 49-63). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, 104(4), 839–862.

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General, 128,* 403–415.

Nairne, J. S. (1996). Short-term/working memory. In E. L. Bjork & R. A. Bjork (Eds.), *Memory. Handbook of perception and cognition* (2nd ed.) (pp. 101-126). San Diego, CA US: Academic Press.

Norman, K.A., Newman, E.L. & Detre, G.J. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review* 114 (4), 887-953.

Peterson, L. R., & Peterson, M. J. (1959). Short term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193-198.

Posner, M. I. (1992). Attention as a cognitive and neural system. *Current Directions in Psychological Science, 1*(1), 11-14.

Power, M. J., Dalgleish, T., Claudio, V., Tata, P., & Kentish, J. (2000). The directed forgetting task: Application to emotionally valent material. *Journal of Affective Disorders, 57,* 147–157.

Raaijmakers, J. G. W. (1979). Retrieval from long-term store: A general theory and mathematical models. Druk: Stichting Studentenpers, Nijmegen.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Roediger, H. L. & Crowder, R. G. (1972). Instructed forgetting: Rehearsal control or retrieval inhibition (repression)? *Cognitive Psychology*, *3*(2), 244-254.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89*(1), 63-77.

Sahakyan, L. (2004). The destructive effects of the "forget" instructions. *Psychonomic Bulletin* & *Review, 11,* 555–559.

Sahakyan, L. & Delaney, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list method paradigm? *Journal of Memory and Language*, *48*(1), 195-206.

Sahakyan, L., & Delaney, P. F. (2005). Directed Forgetting in Incidental Learning and Recognition Testing: Support for a Two-Factor Account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(4), 789-801.

Sahakyan, L., & Goodmon, L. B. (2007). The influence of directional associations on directed forgetting and interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(6), 1035–1049.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1064–1072.

Sahakyan, L., Waldum, E. R., Benjamin, A. S., & Bickett, S. P. (2009). Where is the forgetting with list-method directed forgetting in recognition? *Memory & cognition*, *37*(4), 464.

Sheard, E. D., & MacLeod, C. M. (2005). List method directed forgetting: Return of the selective rehearsal account. In N. Ohta, C. M. MacLeod, & B. Uttl (Eds.), *Dynamic cognitive processes* (pp. 219-248). Tokyo: Springer-Verlag.

Shiffrin, R. M., Ratcliff, R., & Clark, S. E. (1990). List-strength effect: II. Theoretical mechanisms. *Learning, Memory, 16*(2), 179–195.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review, 4,* 145–166.

Shiffrin, R. M., & Steyvers, M. (1998). The effectiveness of retrieval from memory. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 73–95). London: Oxford University Press.

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 460–471.

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: The effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1207–1241.

Waugh, N. C. (1970). On the effective duration of a repeated word. *Journal of Verbal Learning and Verbal Behavior, 9*(5), 587-595. 2

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

Weiner, B. (1968). Motivated forgetting and the study of repression. *Journal of Personality*, *36*, 213-234.

Weiner, B., & Reed, H. (1969). Effects of the instructional sets to remember and to forget on short-term retention: Studies of rehearsal control and retrieval inhibition (repression). *Journal of Experimental Psychology, 79*(2), 226–232.

Whitten, W. B. (1978). Output interference and long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory, 4*(6), 685-692.

Wickelgren, W. A. (1964). Size of rehearsal group and short-term memory. *Journal of Experimental Psychology, 68*(4), 413-419.

Widaman, K. F. (2008). Integrative perspectives on cognitive aging: Measurement and modeling with mixtures of psychological and biological variables. In S. M. Hofer & D. F. Alwin (Eds.) *Handbook of cognitive aging: Interdisciplinary perspectives*, (pp. 50-68). Los Angeles: Sage Publications.

Winograd, E. (1968). List differentiation, recall, and category similarity. *Journal of Experimental Psychology*, *78*(3), 510-515.

Woodward, A.E. & Bjork, R.A. (1971). Forgetting and remembering in free recall: Intentional and unintentional. *Journal of Experimental Psychology, 89,* 109-116.

Woodward, A.E., Bjork, R.A., & Jongeward, R.H. (1973). Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavior, 12*, 608-617.

Xu, J., & Malmberg, K. J. (2007). Modeling the effects of verbal and nonverbal pair strength on associative recognition. *Memory & Cognition, 35*(3), 526-544.