AUTOMATIC AND CONTROLLED EFFECTS ON MEMORY ORGANIZATION

By

Abigail M. D. Mundorf

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology—Doctor of Philosophy

2024

ABSTRACT

The temporal contiguity effect (TCE) is the finding that recalling one event often leads to recalling another event originally experienced nearby in time. Theories of episodic memory differ in their explanations of the TCE, attributing the TCE either primarily to fundamental, automatic memory mechanisms or strategic control processes. The TCE has been well-replicated even under incidental encoding conditions, supporting accounts based on automatic mechanisms. However, the size and shape of the TCE varies across tasks and individuals, suggesting the TCE is also affected by strategic control processes. This dissertation tests the predictions of retrieved context models, which propose the TCE results from the automatic mechanisms of association formation during encoding and context reinstatement at retrieval, and accounts that emphasize the role of control processes in determining recall organization. Four experiments were conducted to test the predictions of these two accounts. In Experiment 1, evidence for a TCE was found in an implicit memory test, supporting retrieved context models' prediction that temporal information is not only automatically encoded but also automatically retrieved. In Experiment 2, a deeper encoding task increased both recall and the TCE as predicted by retrieved context models. However, both recall and the TCE were highest with no assigned encoding task, suggesting control processes also play an important role in recall organization. Experiment 3 directly compared the effect of strategic control processes at encoding and retrieval. The TCE was present in all conditions, but retrieval strategies influenced the degree of both temporal and semantic recall organization. Finally, Experiment 4 tested if temporal information guides recall even when other useful associations are available and task-relevant. Participants primarily displayed temporal organization and did not cluster their recalls based on other kinds of associations even they were task-relevant. Two implementations of retrieved context models were fit to data from Experiments 3 and 4 to test potential implementations of control processes for these models and evaluate the models' ability to explain organization along multiple dimensions. Together, these experiments suggest a comprehensive theory of memory must include both automatic and controlled mechanisms and point to the need for further development of an integrated model of memory.

ACKNOWLEDGEMENTS

My journey towards completing this dissertation would not have been possible without the support and mentorship of my advisor, Dr. Karl Healey. His patience, feedback, and encouragement to always consider how to do things right in principle before putting them into practice have shaped my experience as a researcher and given me both the skills and the confidence to venture into the world of research.

I would also like to thank my committee members, Dr. Erik Altmann, Dr. Kimberly Fenn, and Dr. Zach Hambrick for their helpful feedback throughout this process. Several of their questions and comments encouraged me to go in a new direction with an analysis or re-work an explanation to make broader connections that improved the quality of the research included in this dissertation.

Everyone in the Computational Basis of Cognitive Control Lab also deserves thanks for their role in bringing this research to reality. In particular, Mitchell Uitvlugt, a gem of a lab manager, and Linh Lazarus, who started all of this with me, provided lots of encouragement, logic-checking, and expertise throughout my graduate school career. I am also grateful for the funding provided by the National Science Foundation and the Michigan State University Graduate School for providing funding while I worked on these experiments.

Lastly, my family has been an incredible source of support, from instilling a love of learning and curiosity in me at a young age to believing in me as I worked on my dissertation through a pandemic and many major life events. I would not be where I am today without them. Special thanks is due to my husband and daughter, Christian and Magdalena, for being my inspiration.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Episodic memories are characterized by mental time travel, the reinstatement of an event and its surrounding spatiotemporal context. Episodic memories are also thought to be cue-dependent: a specific episodic memory can only be retrieved in response to an associated retrieval cue (Tulving, 1972). Context can act as that retrieval cue, facilitating access to other memories associated with a similar context. For example, reinstating the environmental context (e.g., on land versus underwater; Godden & Baddeley, 1975; Shin et al., 2021; but see Murre, 2021) or mental context (e.g., emotional state; Bower et al., 1978; for a review, see DuBrow et al., 2017) that prevailed during study results in better recall performance than testing in a different context. Context can also influence recall organization, the *order* in which events are remembered. Such patterns can be observed in free recall, where participants study a list of items and then recall the items in whatever order they come to mind. Participants tend to organize recalls based on based on temporal, spatial, and semantic similarity (Bousfield, 1953; Kahana, 1996; Murdock, 1972; Pacheco & Verschure, 2018; Tulving, 1962). That is, after recalling one event, participants are likely to next recall another event that occurred in a similar context, where context can include a representation of both environmental and mental states.

A well-replicated pattern of recall organization is the temporal contiguity effect (TCE), the tendency for one recall to cue recall of another event previously experienced nearby in time (Kahana, 1996; Murdock, 1974; Postman, 1971). In free recall, the TCE is characterized by a higher likelihood of making transitions between items studied nearby in time with a particularly high bias for recalling those items in the forward direction. Examples of a typical TCE are presented in Figure 1.1. Although the size of the effect varies, the TCE has been observed across several laboratory tasks, including recognition and paired associates tasks (Averell et al., 2016; Campbell & Hasher, 2018; Caplan et al., 2006; Davis et al., 2008; Healey & Kahana, 2016; Schwartz et al., 2005) and with a wide variety of stimuli, from lists of unrelated words (Howard & Kahana, 1999; Kahana, 1996) to autobiographical memories (Diamond & Levine, 2020; Moreton & Ward, 2010;

Pathman et al., 2023). Greater temporal organization also predicts higher recall performance (Polyn et al., 2011; Sederberg et al., 2010; Spillers & Unsworth, 2011; at least in intentional encoding of unrelated lists; Healey & Uitvlugt, 2019; Mundorf et al., 2021), suggesting that memory for items and memory for their order are tightly linked.

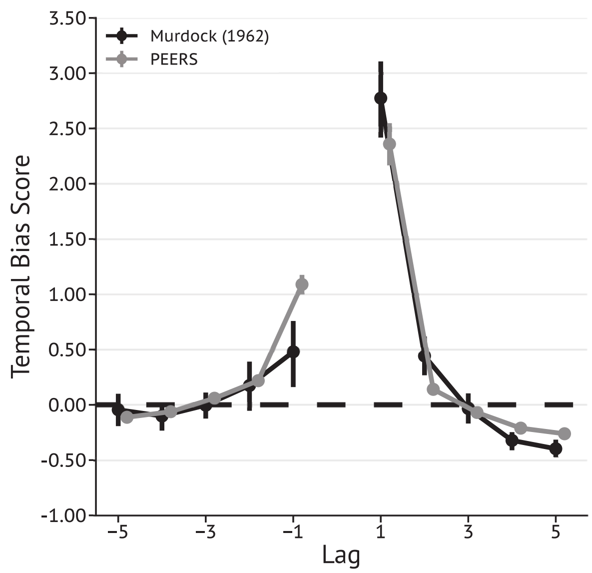

Figure 1.1 Temporal bias curves for two large archival datasets: the list length 15, 2 second presentation rate condition of Murdock (1962) and Experiment 1 of the Penn Electrophysiology of Encoding and Retrieval Study (PEERS). Temporal bias curves represent the probability of making a lag of a given transition relative to what would be expected if the same items were recalled in random order, where lag is the distance between the just-recalled item and the next recall in their original presentation order. For example, recalling the second item in the list followed by the fourth item in the list is a transition of $lag = 4 - 2 = +2$ because recall is advancing two positions forward in the original order of the list. The dashed line indicates a score of zero, which indicates no bias. Error bars indicate bootstrapped 95% confidence intervals. Figure reprinted with permission from Mundorf, Lazarus, Uitvlugt, and Healey (2021).

The ubiquity of the TCE has led theories of episodic memory to include explicit TCE-generating mechanisms (e.g., Davelaar et al., 2005; Farrell, 2012; Howard & Kahana, 2002; Howard et al., 2015; Lehman & Malmberg, 2013). However, the relevance of the TCE for such theories is dependent on the degree to which the TCE is a result of automatic versus controlled memory processes. That is, does temporal organization arise through mechanisms central to the formation and retrieval of episodic memories, or does the TCE merely reflect intentional, task-specific strategies that are engaged only when temporal order is attended to?

Distinguishing Automatic and Controlled Memory Mechanisms

If the TCE arises from fundamental automatic memory mechanisms, then examining patterns of temporal organization may reveal important insights into the human memory system as a whole. In that case, the purpose of researching this effect is to develop theories with mechanisms that explain both recall and the TCE as a result of the same underlying mechanisms. For example, the positive correlation between the TCE and recall may indicate that memory for temporal order is inherently tied to the formation and retrieval of episodic memories. In contrast, if the TCE is entirely a result of strategic control processes, the effect would reveal more about individual differences in adapting to specific tasks than about fundamental memory mechanisms. Perhaps individuals who display a greater TCE also recall more words merely because they adopt successful strategies. Research on a TCE generated by strategic control processes would be useful for understanding this effect for its own sake but would be less important for developing memory theory as a whole. A third possibility is that the TCE occurs through a combination of automatic and controlled processes, and episodic memory theories should include both kinds of mechanisms. In any case, determining the automatic and controlled causes of temporal organization is critical for memory theory development.

In theories of attention, a major difference between automatic and controlled processes is that controlled processes require the intentional use of some kind of limited cognitive resource, while automatic processes do not (e.g., Kahneman, 1973; Treisman & Gelade, 1980). Hasher and Zacks (1979) applied this basic distinction to episodic memory, proposing a framework for differentiating memory phenomena that are due to automatic versus controlled processes. According to their framework, automatic memory processes 1) do not require attention or awareness, 2) do not interfere with other memory processes, and 3) are consistent across individual and task differences. In contrast, controlled processes require effort and intention, interfere with other effortful processes, and can vary widely across individuals and situations. This framework has had a substantial influence in the field of cognitive aging, where it has been used to suggest deficits in controlled processes may be responsible for aging effects in episodic memory (Hasher & Zacks, 1988; Jennings

& Jacoby, 1993; Old & Naveh-Benjamin, 2008; Reuter-Lorenz & Park, 2010). Importantly, these criteria can be applied to any memory phenomenon to determine if it is a result of automatic or controlled processes, including the TCE.

Evidence for Automatic and Controlled Effects on the TCE

If the TCE is due to automatic mechanisms, then temporal organization should occur even in the absence of intentional study, not interfere with other processes, and be consistent across individuals and situations. If the TCE is instead a result of controlled processes, then the TCE should be eliminated when participants are not intentionally trying to encode or retrieve temporal information, be reduced by other interfering processes, and occur only for some individuals and in some situations. The following sections will review research on the TCE and utilize each of Hasher and Zacks's (1979) three criteria to determine if the TCE is a result of automatic or controlled processes. These findings will then be discussed in light of their implications for two classes of memory theories.

1) Does the TCE Occur in the Absence of Intentional Study?

A clear test of whether temporal information is encoded automatically is to examine if a TCE occurs under conditions of incidental encoding, when participants are unaware that their memory will be tested and therefore have no reason to engage in strategic control processes during encoding. If temporal information is learned exclusively through controlled processes, then the TCE should be abolished by incidental encoding. Recent work has found a TCE even when participants were unaware that their memory would later be tested. In Mundorf et al. (2021), I tested participants' memory for a single word list under either intentional or incidental encoding. Participants in the incidental conditions were provided a cover story to prevent them from suspecting a later memory test while participants in the intentional conditions knew their memory would be tested. Regardless of encoding intentionality, participants displayed a clear TCE, as displayed in Figure 1.2 (see also Healey, 2018). Similar results have been obtained with incidental encoding of more complex materials, such as news stories and autobiographical memories (Diamond & Levine, 2020; Moreton & Ward, 2010; Pathman et al., 2023; Uitvlugt & Healey, 2019). Based on Hasher and Zacks's

(1979) criteria, these results strongly support the claim that temporal information is automatically encoded.

However, there is an effect of encoding intentionality on the *size* of the TCE, suggesting that the degree to which temporal information is encoded may rely somewhat on controlled processes. The TCE is greatly reduced under incidental compared to intentional encoding (Healey, 2018; Mundorf et al., 2021; Nairne et al., 2017). And while intentionally-encoding participants show a bias for recalling items in *forward* order, incidental encoding leads to an equally high bias for making transitions to any item studied nearby in time, regardless of if it was previously studied before or after the just-recalled item (see Figure 1.2; Mundorf et al., 2021). One potential explanation for these findings is that the forward bias is a result of intentional encoding strategies like linking the words together to form a story (Bouffard et al., 2018). When participants are unable to engage in those strategies, the forward bias disappears. This difference is consistent with the possibility that temporal information is automatically encoded to some extent but that order-based encoding strategies may also contribute to the TCE.

Is Temporal Information Retrieved in the Absence of Intentional Retrieval? Even though temporal information is automatically encoded to some degree, it is unclear if that same information is also automatically retrieved. In free recall of unrelated word lists, adopting an order-based recall strategy is highly effective (Sederberg et al., 2010; Spillers & Unsworth, 2011). Therefore, the TCE in free recall could be due to participants intentionally retrieving items in the order they were originally studied rather than to automatic mechanisms. If temporal information *is* automatically retrieved, then temporal information should guide memory search even when order-based strategies are irrelevant or hurt performance.

There is some evidence that temporal information is automatically retrieved from paired associates recall, where participants recall one pair member given the other member as a cue. In a typical paired associates task, remembering extra-pair items presented right before or after a given pair can be detrimental to memory performance by increasing the likelihood of making an intrusion. Therefore, there is no reason for participants to intentionally remember the order of the

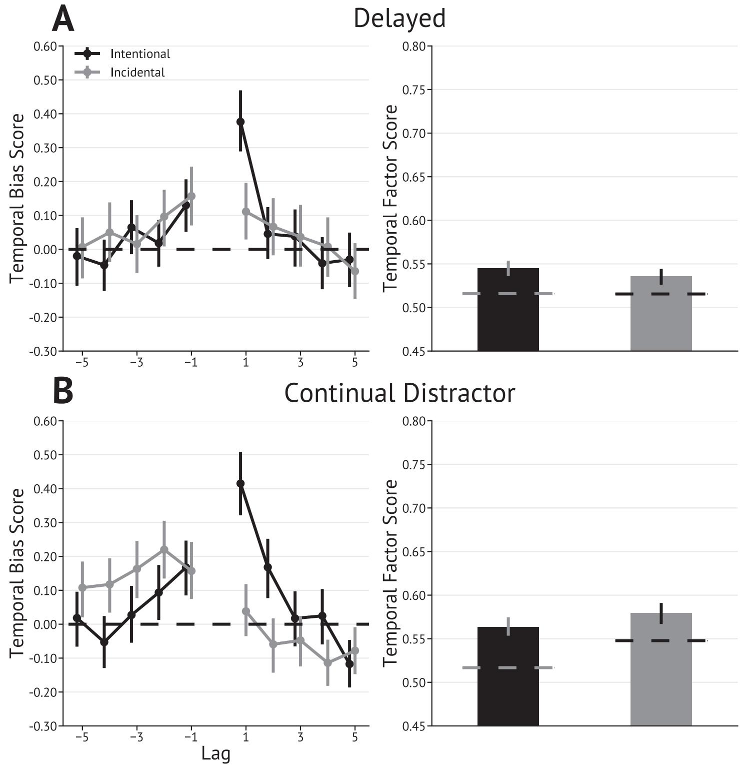

Figure 1.2 Temporal bias curves and temporal factor scores for the four conditions in Mundorf et al. (2021). Participants either incidentally or intentionally encoded items in either A) delayed free recall, where a filled delay intervened between encoding and recall or B) continual distractor free recall, where a filled delay followed each item. Temporal bias curves (left) represent the bias for making a transition of a given lag beyond what would be expected if the items were recalled in random order. The dashed line indicates a score of zero, which indicates no bias. Temporal factor scores (right) are a single-number measure of the TCE where scores above chance (dashed line; calculated individually for each condition) indicate a significant TCE. Error bars are bootstrapped 95% confidence intervals. Figure reprinted with permission from Mundorf, Lazarus, Uitvlugt, and Healey (2021).

pairs in the list. Nonetheless, Davis et al. (2008) found that intrusions were more likely to come from pairs originally studied nearby in time to the target pair, as would be expected if temporal information were automatically retrieved (see also Caplan et al., 2006).

However, findings here are also mixed. Osth and Fox (2019) found no evidence of increased false alarms for lures studied nearby in time to the cue in a paired associates *recognition* task. They concluded that temporal context information can be suppressed at retrieval, or even at encoding, if temporal order is not task-relevant. Why might these studies have come to such different conclusions? One possibility is that temporal context information is automatically retrieved, but some participants are able to ignore it more effectively. Campbell et al. (2014) found that only older adults, who likely have less ability to inhibit the retrieval of temporal context than younger adults, displayed a TCE in paired associates recognition. The undergraduate students participating in the Osth and Fox (2019) study may have simply been more effective at engaging control processes to suppress temporal context information. These results leave open the possibility that temporal information is automatically retrieved, but the answer is still unclear. One limitation of these studies is that in paired associates tasks, participants are still engaging in intentional retrieval, even if they are not intentionally recalling the study order. It remains to be seen if temporal information is remembered *whenever* an item is retrieved, even when retrieval itself is unintentional.

2) Does the TCE Interfere with Other Processes?

It is unclear if other cognitive processes interfere with the mechanisms responsible for the TCE. According to Hasher and Zacks's (1979) framework, automatic processes should not compete with controlled processes for limited attentional resources. Therefore, if temporal information is automatically encoded, dividing attention between memory processes and some other task which requires cognitive control should not interfere with the TCE. Murphy and Castel (2021) tested this prediction and compared recall organization for participants who completed a simple divided attention task during encoding (making judgements about tones played during study) to those who encoded the items with their full attention. Although recall was higher for the full attention condition, the TCE was unaffected by divided attention during encoding, regardless of the difficulty of the secondary judgement task. They concluded that encoding temporal information did not compete with the tone task for limited cognitive resources, and temporal information is automatically encoded.

At the same time, assigning a *task-relevant* encoding task may reduce the TCE. Task-relevant encoding tasks are often assigned in free recall either to boost recall performance or to ensure participants are paying attention. These encoding tasks differ from a standard divided attention task in that they encourage participants to engage in additional processing of the stimuli they are supposed to be encoding rather than dividing their focus between study and a separate task.

However, if the TCE is a result of order-based strategies, assigning any task might interfere with those strategies and reduce the TCE. Long and Kahana (2017) found that assigning participants a semantic processing task during encoding ("Does this word refer to something living or nonliving?") not only reduced both recall and the TCE but also interfered with neural activity associated with the formation of new episodic associations.

The difference in findings between Murphy and Castel (2021), who found no effect of divided attention on the TCE, and Long and Kahana (2017), who found that assigning an encoding task reduced the TCE, could be due to differences in the encoding tasks. It is possible that the divided attention task used by Murphy and Castel (2021) was too simple to interfere with any strategic control processes that may have contributed to the TCE, while Long and Kahana's (2017) semantic processing task was difficult enough to interfere with participants' use of order-based encoding strategies. Alternatively, the materials involved in the additional task may matter. Murphy and Castel's (2021) divided attention task required participants to make judgements about stimuli unrelated to the main encoding task. Therefore, they may have not interfered directly with encoding processes. Semantic processing of the study items, on the other hand, may have more directly interfered with the encoding of temporal context information in Long and Kahana's (2017) experiment by biasing participants to focus on meaningful relationships between items rather than order.

Overall, these findings provide mixed evidence for the TCE being a result of automatic mechanisms; some tasks may interfere with encoding temporal information, while others may not. This lack of clarity points to a critical gap in the current literature. Future work could address the effect of an additional encoding task on the TCE by varying encoding task difficulty with particular attention to how an encoding task affects not only the TCE, but also the relationship between recall and contiguity. For example, some encoding tasks, such as deeply processing items for their semantic meanings, have been found to increase recall performance (Craik & Tulving, 1975). If temporal information is automatically encoded whenever a new memory is formed, any task that increases recall should also increase the TCE. In contrast, any assigned encoding task would likely interfere with any controlled processes that contribute to the TCE, so assigning a semantic processing task during encoding would reduce a TCE due to strategic control processes.

3) Is the TCE Consistent Across Individuals and Situations?

Individual Differences in Temporal Contiguity. Because automatic memory processes form the basis of all memory formation and retrieval, Hasher and Zacks (1979) propose there should be very little variation in automatic processes across individuals. In contrast, controlled processes rely on the availability of attentional resources, so anything that reduces available attentional resources or cognitive control should also reduce effects caused by controlled processes.

We can test the TCE against these criteria. The effect of temporal context is robust across individuals. Healey and Kahana (2014) examined recalls across several sessions for 126 younger adults and found that, depending on how the TCE was measured, between 96% and 100% of participants displayed a TCE in their recalls. The TCE has also been observed across the lifespan, from elementary school students (Lehmann & Hasselhorn, 2010, 2012; Pathman et al., 2023) to older adults (Healey & Kahana, 2016; Howard et al., 2006; Kahana et al., 2002; Wahlheim & Huff, 2015), as would be expected if the TCE were due to automatic mechanisms. And within individuals the TCE is unaffected by fluctuations in attention throughout a list (Jayakumar et al., 2023).

However, as noted by Healey et al. (2019), the size and shape of the TCE varies substantially across individuals. Participants' ability for cognitive control may be a key source of this variation. Individuals with higher scores on tests of IQ and working memory, tests considered to measure cognitive control, display greater temporal organization (Healey & Kahana, 2014; Spillers & Unsworth, 2011). And although the TCE is present in both young children and older adults, the size of the TCE varies with age. In young children, the TCE increases with age (Lehmann & Hasselhorn, 2010; Pathman et al., 2023), possibly due to older children developing more effective encoding and retrieval strategies (Lehmann & Hasselhorn, 2012). Temporal contiguity is also greater for younger adults compared to older adults and for healthy older adults compared to older adults who are later diagnosed with cognitive decline (Diamond & Levine, 2020; Healey & Kahana, 2016; Howard et al., 2006; Kahana et al., 2002; Talamonti et al., 2021). Children and older adults may both have an impaired ability to engage in strategic control processes that promote recalling items in order, resulting in a smaller TCE than for healthy younger adults (Lehmann & Hasselhorn, 2012; Talamonti et al., 2021). In addition, the TCE is reduced by a number of clinical and personality variables related to reduced cognitive control, including high trait worry (Pajkossy et al., 2017), schizophrenia and high risk for future schizophrenia diagnosis (İmamoğlu et al., 2022; Polyn et al., 2015; Sahakyan & Kwapil, 2018), and experiencing an initial episode of psychosis (Murty et al., 2018; for an example of a clinical condition *increasing* the TCE, see Gibson et al., 2019). Overall, individuals with a reduced capacity for cognitive control display reduced temporal organization. These differences suggest that strategic control processes likely play an important role in generating the TCE.

Task Differences in Temporal Contiguity. According to Hasher and Zacks's (1979) framework, automatic memory processes should occur across different situations because whereas controlled processes are developed to improve performance on a specific task, automatic processes rely on mechanisms that operate *whenever* memories are formed or retrieved. Therefore, a TCE that appears only in some tasks is likely to be a result of strategic control processes, but a TCE that is observed regardless of task parameters is consistent with being caused by automatic memory mechanisms.

Task Manipulations. The TCE has been found across a wide variety of different task manipulations and with stimuli of varying complexity, supporting claims that the effect is due to automatic mechanisms. Evidence for temporal organization has been found not only in free recall, but also in paired associates tasks (Campbell & Hasher, 2018; Campbell et al., 2014; Caplan et al., 2006; Davis et al., 2008; but see Osth & Fox, 2019) and recognition (Averell et al., 2016; Healey & Kahana, 2016; Sadeh et al., 2015; Schwartz et al., 2005; but see Bradley & Glenberg, 1983). The TCE also occurs even at very rapid presentation rates, when it is unlikely participants are able to engage in any encoding strategies (up to 8 words per second; Toro-Serey et al., 2019). In addition, temporal contiguity has been observed when to-be-remembered items are separated by intervening distractor events (as in continual distractor free recall; Bhatarah et al., 2006; Howard & Kahana, 1999; Mundorf et al., 2021) and when items are experienced several seconds, hours, or even days apart (Mack et al., 2017; Moreton & Ward, 2010; Nguyen & McDaniel, 2015; Uitvlugt & Healey, 2019). It is unlikely that participants would be able to engage in order-based encoding strategies across such long delays.

Although the TCE is largely robust to differences in task parameters, some manipulations can reduce the size of the effect, particularly those manipulations that change the usefulness of temporal information. For example, the TCE is reduced in longer lists where temporal order information may be less useful in guiding retrieval (Cortis et al., 2015; Healey et al., 2019). The TCE is moderated by some task manipulations, suggesting that control processes play at least some role in producing temporal recall organization.

Stimuli Characteristics. Hasher and Zacks (1979) also predict automatic memory processes should operate regardless of stimulus complexity. In contrast, controlled processes adapt to the stimulus features, so memory effects due to controlled processes should vary across stimuli. The TCE has been observed with stimuli of varying complexity, including images, lists with semantic structure, and even nonverbal stimuli (Cortis et al., 2015; Healey & Uitvlugt, 2019; Nguyen & McDaniel, 2015; Polyn et al., 2011). The TCE has also been found with naturalistic stimuli (Diamond & Levine, 2020; Moreton & Ward, 2010; Pathman et al., 2023; Uitvlugt & Healey, 2019). For example, Diamond and Levine (2020) examined temporal contiguity for younger and older adults' recalls of a museum tour. Both younger and older adults displayed substantial temporal clustering and forward asymmetry; that is, after recalling one event from the tour, they were much more likely to recall another event that occurred nearby in time, particularly if it was the very next event that occurred. Autobiographical memories such as those recalled in this experiment include a complex web of not only temporal context associations, but also semantic, spatial, and emotional associations. The finding of a TCE even with these rich autobiographical memories provides strong support for the claim that temporal information is automatically processed.

However, when other kinds of associations are present and temporal order is not related to similarity along other associative dimensions, the TCE is reduced (Healey & Uitvlugt, 2019; Polyn et al., 2011). A shift in organization strategies may be responsible for the reduced TCE. For example, when items are semantically related, focusing on semantic relationships between words is an effective recall strategy (Mandler, 1967; Tulving & Pearlstone, 1966). Participants may opt to organize their recalls based on semantic similarity, reducing the TCE if semantically related items are not presented in adjacent serial positions. When lists are designed such that semantically related items are not presented in temporal order, the TCE is substantially reduced (Healey & Uitvlugt, 2019; Polyn et al., 2011), and temporal organization no longer predicts recall success (Uitvlugt & Healey, 2020). Instead, recall is correlated with semantic organization (Healey & Uitvlugt, 2019). The TCE may even be eliminated by combining multiple factors that reduce the size of the TCE, such as testing participants on a long list with semantic structure (Hong et al., 2023). These variations challenge the claim that the TCE is generated through automatic processes. Based on Hasher and Zacks's (1979) framework, if the TCE is entirely due to automatic processes, it should occur even in situations where participants decide to adopt a different strategy.

Healey and Uitvlugt (2019) not only examined the effect of the task materials, but they also directly manipulated recall strategies by instructing participants to adopt either a meaning-focus, temporal-focus, or a free recall strategy while recalling lists that contained several semantic clusters (related lists) or were composed entirely of unrelated words (unrelated lists). The TCE was not only reduced but was actually eliminated in related lists when participants adopted a meaning-focus strategy. This suggests that the size of the TCE is greatly influenced by intentional strategies, and participants have the capacity to ignore temporal order information if it is not relevant to their current goals. At the same time, there was also evidence that temporal information was automatically encoded even if it was not intentionally used to guide recall. *Within* a semantic cluster, participants were more likely to transition between cluster members originally studied closer in time. That is, there was a within-cluster TCE, even though cluster members were not studied in adjacent serial positions (see also Polyn et al., 2011). This study provides strong evidence for both the automatic and controlled explanations of the TCE. Participants' recall organization is under their conscious control, but temporal information is also automatically encoded and may influence recall even when participants are trying to ignore it.

Interim Summary

In summary, there is evidence that temporal context effects in memory are automatic to some extent. Temporal information is available even when participants are prevented from using order-based encoding strategies, and assigning an additional task during encoding does not eliminate the TCE although it may reduce the TCE if the additional task encourages participants to adopt an alternate method of retrieval organization. Nearly all individuals show a bias for recalling items experienced nearby in time together even when they are not intentionally studying or when recalling in temporal order interferes with performance, regardless of task manipulations or stimuli characteristics. This is consistent with Hasher and Zacks's (1979) criteria for information that is processed automatically in memory.

Yet, the size and shape of the TCE vary greatly across individuals and situations. Notably, the TCE is reduced when participants have a reduced capacity for cognitive control, suggesting a role for control processes. When participants are not expecting a memory test and therefore are unlikely to be engaging in order-based strategies, the TCE is greatly reduced compared to intentional study. Similarly, cognitive aging and clinical conditions associated with impaired cognitive control lead to a reduced TCE relative to healthy younger adults. And when alternate encoding or retrieval strategies are more advantageous, the TCE is reduced or even eliminated. Therefore, the evidence strongly suggests that both automatic and controlled processes are responsible for generating the TCE.

Implications of the TCE for Episodic Memory Theories

The influence of both automatic and controlled processes on the TCE make it a useful tool for testing different theories of memory which explain the TCE through very different mechanisms. Although most theories predict a TCE in free recall, some assume the TCE is a result of control processes alone, while others attribute the TCE to automatic memory mechanisms that are fundamental to the formation and retrieval of episodic memories. These theories differ in their predictions for the specific circumstances under which a TCE should be observed, so we can test these theories by examining how well they account for the TCE across individuals and tasks.

Control Processes Accounts of the TCE

Accounts of recall organization based on control processes assume that the TCE is a result of intentional strategies during encoding and retrieval and temporal information is *not* encoded automatically (e.g., Hintzman, 2016). Control processes refer to intentional memory strategies that may be interfered with by other intentional processes. For example, participants may engage in rehearsal during encoding or develop a retrieval plan to recall items in the exact order they were originally studied. The TCE has been primarily studied in free recall tasks where participants know they will be tested and intentionally search memory for as many items as possible. These participants have every opportunity to engage in strategic control processes at encoding or retrieval, even if they are not explicitly instructed to do so. As a result, the patterns of recall organization observed in the lab could be due to control processes alone (Hintzman, 2011).

Control processes accounts of the TCE are based on the assumption that participants intentionally focus on temporal order during encoding and retrieval to help guide their memory search. In free recall, participants do often spontaneously adopt strategies that encourage temporal organization like connecting items together to form a story or rehearsing items in order (Delaney & Knowles, 2005; Hintzman, 2016; Unsworth, 2016). These order-based strategies may directly contribute to the TCE by encouraging recalling the items in forward order (enhancing the forward asymmetry that is characteristic of the TCE; Bouffard et al., 2018; Unsworth et al., 2019). These strategies are also highly effective. In lists of unrelated words, the TCE is positively correlated with recall (Polyn et al., 2011; Sederberg et al., 2010; Spillers & Unsworth, 2011). Participants who do well on free recall could simply be those who quickly identify and adopt a useful strategy. The TCE also increases with task experience (Healey et al., 2019). As participants learn what kinds of strategies are most effective, their recalls become more organized, and recall tends to improve. In this case, increased temporal contiguity may reflect participants developing a task-specific strategy for free recall that may not generalize to other tasks.

According to a control processes account, anything that reduces participants' ability to effectively engage in order-based encoding or retrieval strategies should reduce the TCE. Consistent with this prediction, the TCE is reduced among individuals with a reduced capacity for cognitive control, such as older adults and young children (Healey & Kahana, 2016; Howard et al., 2006; Lehmann & Hasselhorn, 2010; Pathman et al., 2023). In addition, a control processes account predicts that the TCE should be present only so long as temporal information is useful. When other kinds of associations are present, participants cluster their recalls based on semantic similarity (Healey & Uitvlugt, 2019; Polyn et al., 2011; Tulving & Pearlstone, 1966), spatial location (Bouffard et al., 2018; Curiel & Radvansky, 1998; Gibson et al., 2021; Miller et al., 2013), emotional valence (Long et al., 2015; Siddiqui & Unsworth, 2011; Talmi et al., 2019), level of associated reward (Murphy & Castel, 2021; Stefanidi et al., 2018), task rules (Polyn et al., 2009a, 2009b), or even surface features like presentation modality (Hintzman et al., 1972; Nilsson, 1974). When alternate retrieval strategies are encouraged, the TCE is often reduced and temporal contiguity is no longer correlated with overall recall (Clark & Bruno, 2021; Healey & Uitvlugt, 2019; Hong et al., 2023; Pacheco & Verschure, 2018; Polyn et al., 2011; but see Talamonti et al., 2020), suggesting that temporal information is either not encoded or not used to guide retrieval when it is less useful. A control processes account naturally predicts this shift in organization by assuming that participants adapt strategies to task conditions.

All of this evidence points to a role for control processes in producing temporal contiguity. Yet, a strict interpretation of a control processes account has difficulty explaining the robust TCE observed in situations where participants are unlikely to be intentionally linking items together. For example, this account is inconsistent with the TCE found under incidental encoding (Diamond & Levine, 2020; Healey, 2018; Moreton & Ward, 2010; Mundorf et al., 2021; Uitvlugt & Healey, 2019) and when temporal information is not useful for retrieval (Davis et al., 2008; Healey & Uitvlugt, 2019; Polyn et al., 2011). Similarly, an account that attributes the TCE entirely to strategic control processes would have difficulty explaining the temporal contiguity observed among items that are separated by other events or presented hours or even days apart (Howard & Kahana, 1999; Mack et al., 2017; Moreton & Ward, 2010; Mundorf et al., 2021; Uitvlugt & Healey, 2019), when it is unlikely that participants would be rehearsing or intentionally connecting the items.

Therefore, a major challenge for accounts of the TCE based on control processes is that they fail to account for the TCE observed when participants are prevented from engaging in order-based strategies or encouraged to use a different strategy. Another challenge is that such accounts rarely include well-specified mechanisms, making them useful for *describing* the effects of strategic control processes but less useful in *explaining* the underlying mechanisms for how control processes affect memory formation and retrieval. These accounts are consistent with the finding that, for example, linking items together to form a story both improves recall and increases the TCE. Yet, the mechanisms responsible for the strategic control processes that affect the TCE, recall, and the relationship between them are not clearly defined in accounts based on control processes.

Retrieved Context Models

An alternate explanation of the TCE that relies on automatic, rather than controlled, mechanisms is provided by retrieved context models, a family of computational models that emphasize the role of temporal context at both encoding and retrieval (for recent implementations, see Healey & Wahlheim, 2023; Howard & Kahana, 1999; Howard et al., 2015; Lohnas et al., 2015; Polyn et al., 2009a). Computational models of memory in general require cognitive mechanisms to be clearly specified so they can be implemented in the form of mathematical equations. Retrieved context models in particular are well-suited to explain the ubiquity of the TCE because they attribute both memory performance and the TCE to the same automatic contextual dynamics.

Retrieved context models have been applied to explain not only temporal contiguity but also a variety of other memory phenomena, including age-related memory change (Healey & Kahana, 2016, 2020; Howard et al., 2006; Wahlheim & Huff, 2015), amnesia (Palombo et al., 2019; Sederberg et al., 2008), individual differences in cognitive performance (Healey & Kahana, 2014), spacing effects (Kahana & Howard, 2005; Siegel & Kahana, 2014), emotional memory (Long et al., 2015; Talmi et al., 2019), consolidation (Sederberg et al., 2011), interference (Lohnas et al., 2015), retrieval practice effects (Hong, Polyn, & Fazio, 2019; Karpicke et al., 2014; Whiffen & Karpicke, 2017), scene exploration (Kragel & Voss, 2021), and event segmentation (Ezzyat & Davachi, 2014; Pu et al., 2022; Sahakyan & Smith, 2014). The same mechanisms have also been implemented in models that incorporate other task features, including semantic relationships among studied items (Polyn et al., 2009a) and response latencies (Sederberg et al., 2008).

The automatic mechanisms of context drift, association, and reinstatement are the basis of these models' ability to account for the TCE and other memory phenomena. An example of the encoding period, as described in these models, is presented in Figure 1.3. Under these models, new episodic memories are created when an association is formed between a representation of the item being studied (represented on the feature layer in Figure 1.3) and the current state of temporal context (represented on the context layer in Figure 1.3). Here, temporal context represents whatever is active in mind at a given point in time. As the current item is processed, it brings to mind its pre-existing associations which are then incorporated into the current state of context, causing context to change, or drift, as each item is studied. In this way, context drift is directly driven by the items themselves. When the next item is studied, it automatically forms a new association with this updated context and then brings to mind its own existing associations, causing context to drift further. When the next item is studied, the previous context representation is not completely erased; context is a blend of the previous context and the new context. As a result, items experienced relatively closer together in time are associated with more similar states of context.

At retrieval, context is used as a cue. The current context serves as a better cue for items originally associated with a more similar state of context; therefore, items experienced in a similar context to the current context are more likely to be recalled next. Once an item is recalled, it brings back to mind both its pre-existing associations and its associated context from encoding. These associations are then incorporated into the current state of context which is used as a retrieval cue for the next item. Because items experienced closer together in time are associated with more similar states of context, the reinstated context is a better cue for other items experienced nearby in the list. Therefore, the models naturally predict a TCE.

Importantly, retrieved context models assume memory for temporal order is a result of automatic mechanisms: context association and drift during encoding and the reinstatement of those associations when an item is retrieved. Therefore, they naturally predict a TCE under almost any

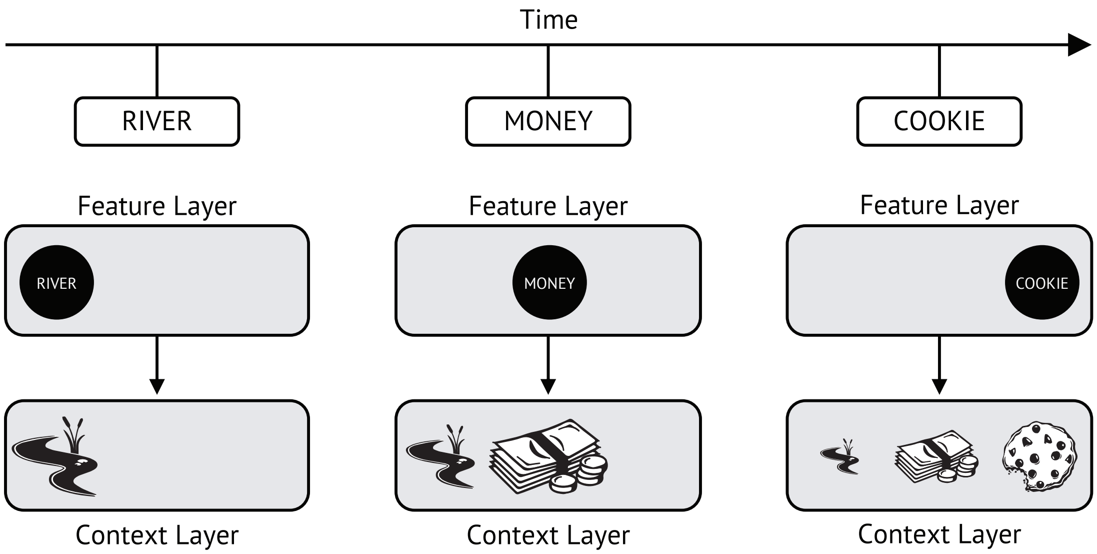

Figure 1.3 Visual representation of the encoding period under retrieved context models. The identities of items are represented on the feature layer, with one node for each item. Each item has a corresponding node on the context layer which represents that item's contextual associates. When the first item, *river*, is presented, its node becomes active on the feature layer, which then activates the *river*-related node on the context layer. When the next item, *money*, is presented, its representation on the feature layer becomes active, completely replacing the *river* activation on the feature layer. Next, the *money*-related context becomes active on the context layer but does not entirely replace the *river* context. Instead, the *river* context fades so that the current mental context is a blend of *river* and *money* contexts in which more recent items are more highly activated. The mental context serves as a recency weighted record of the past. This process repeats when *cookie* is presented: its node on the feature layer becomes active, and its context representation is activated on the context layer, blending with other elements on the context layer to create a new *river-money-cookie* context.

circumstance, including under conditions of incidental encoding, a prediction supported by recent work (Healey, 2018; Mundorf et al., 2021). The framework is also consistent with the TCE found at various timescales, including when items are separated by several minutes, hours, or days (Mack et al., 2017; Moreton & Ward, 2010; Uitvlugt & Healey, 2019) because under these models, items experienced closer in time are associated with *relatively* more similar states of context than those experienced farther apart in time, regardless of if items were experienced 10 seconds or 10 minutes apart. These models are also supported by neuroimaging work. Patterns of neural activity change gradually during encoding, consistent with the context drift mechanism described by these models, and successful retrieval is accompanied by the reinstatement of the pattern of neural activity that prevailed when the item was originally studied, consistent with the retrieved context model mechanism of context reinstatement (Chan et al., 2017; Deuker et al., 2016; Ezzyat & Davachi, 2014; Howard et al., 2012; Manning et al., 2011; Manns et al., 2007; Sederberg, Schulze-Bonhage, Madsen, Bromfield, Litt, et al., 2007).

A strength of these models is that they provide a well-specified, testable account of both memory performance and recall organization. Because the same mechanisms underlie both successful memory formation and retrieval as well as temporal contiguity, the models naturally predict a TCE under almost any circumstance (see Mundorf et al., 2021, for a discussion of cases where these models could predict a null TCE). They also allow for different degrees of temporal contiguity through variations in model parameters. For example, differences in temporal contiguity based on different encoding instructions may be explained by different rates of context drift during encoding. Faster context drift during encoding could increase the TCE by making only items experienced very close in time good cues for one another (as suggested by Healey & Kahana, 2016).

However, a weakness of these models is that they do not specify *how* these differences in parameters are determined, making it difficult to make predictions about when, why, or how recall organization might vary across different situations. In model simulations, parameter values are either selected *a priori* by the researchers based on theory or determined using a search algorithm to maximize model fit. But in actual participants, there must be cognitive control mechanisms that interact with the memory system to determine the degree to which temporal information is encoded and later influences recall, perhaps by tuning parameters of the memory system. These cognitive control mechanisms are missing from these models.

Integrating Control Processes and Retrieved Context Models

Accounts of the TCE based on control processes and retrieved context models both provide a partial explanation of recall organization. Although these two accounts have often been framed in opposition to one another (e.g., Healey et al., 2019; Hintzman, 2016), integrating automatic contextual dynamics with mechanisms for control processes would provide the most comprehensive explanation of existing data. Accounts based on control processes are consistent with the variation in the TCE across individuals and situations but are unable to explain temporal organization under conditions where participants are prevented from engaging in order-based strategies. Retrieved context models provide a well-specified explanation of the automatic mechanisms underlying recall dynamics. However, a representation of strategic control processes in these models has not been fully developed, making it difficult to make *a priori* predictions about the effects of task variations on the TCE.

We can gain some insight into how control processes might be represented in retrieved context models by considering how existing models of cognitive control utilize mental context to represent intentional strategies and how those strategies influence behavior. In connecting these two literatures, context plays a key role. Retrieved context models assume that the formation and retrieval of associations between items and the current state of mental context are fundamental memory processes. Neural and computational models of cognitive control emphasize the role of brain regions such as the prefrontal cortex in maintaining goals as a part of mental context and biasing context representations in other brain regions to align with those goals. The following section will discuss mechanisms from theories of cognitive control and some recent work attempting to integrate these processes into existing memory models.

Theories of Cognitive Control: Neural Mechanisms Underlying Cognitive Control of Memory

Neural theories of cognitive control that address the effects of cognitive control on memory highlight the role of two regions: the hippocampus and the prefrontal cortex (PFC). Both regions are thought to represent and influence temporal context in different ways.

It is well-established that the hippocampus is critical for both encoding and retrieval of episodic memories (Corkin, 1968; Gabrieli, 1998; Scoville & Milner, 1957; Squire et al., 2001), particularly memories for the original context in which an episodic memory was experienced. Recollection memory, which involves a more complete retrieval of the encoding context than judgements based purely on familiarity (Yonelinas et al., 2002), is highly dependent on the hippocampus (Eichenbaum et al., 2007; Viskontas et al., 2009; Yonelinas et al., 2002; Yonelinas et al., 2005). Hippocampal damage also impairs performance on direct tests of context memory, such as asking participants to remember the time, place, or source of a memory (Mastrogiuseppe et al., 2019; McClelland et al., 1995; Navawongse & Eichenbaum, 2013; Nielson et al., 2015; Squire et al., 2001), and greater hippocampal activation during encoding is associated with better memory for both item-item and item-context associations on a later test (Staresina & Davachi, 2009). This has led many to suggest that an important function of the hippocampus is automatically binding items to their episodic context (Ekstrom & Ranganath, 2018; Moscovitch, 1992; O'Reilly et al., 1999; Polyn & Kahana, 2008).

The PFC is generally considered to be the center of cognitive control. Extensive work with both humans and animals has demonstrated the PFC is critical for many executive functions which require control, including decision-making, determining appropriate responses, making inferences, and maintaining information and task goals in the face of interference (Eichenbaum, 2017a; Miller & Cohen, 2001; O'Reilly et al., 2010; Picton et al., 2007; Preston & Eichenbaum, 2013; Unsworth & Engle, 2007; for a discussion on how cognitive control networks are distributed across the brain, see Camilleri et al., 2018; Duncan, 2010; O'Reilly et al., 1999). The role of the PFC in influencing memory can be conceptualized in retrieved context models in terms of control processes' influence on temporal context. Working from a background in retrieved context models, Polyn and Kahana (2008) proposed that activity in the dorsolateral PFC maintains current task goals as a part of mental context. This suggestion is consistent with many existing models of cognitive control unrelated to retrieved context models, where the PFC is assumed to be specialized for the representation of task goals as a part of the current state of mental activity (context) and flexibly adjusting that activity when task goals change (see Duncan, 2010; Hazy et al., 2006; Miller & Cohen, 2001; O'Reilly et al., 1999; Wagner, 2002). Polyn and Kahana (2008) also assume that the hippocampus is primarily responsible for forming episodic associations between items and context, while items themselves are represented elsewhere in the temporal lobe. Communication between the PFC and hippocampus occurs at both encoding and retrieval, allowing the hippocampus to transmit information about the current context to the PFC and for context representations in the PFC to bias retrieval from the hippocampus. A visual representation of this framework is presented in Figure 1.4.

Under this framework, the PFC influences episodic memory in two ways: 1) the PFC maintains task goals as a part of context and then 2) influences processing in the hippocampus and other related brain regions to affect memory processing at both encoding and retrieval. The next sections will expand on each of these mechanisms, discuss evidence for their existence, and consider they fit into not only theories of cognitive control but also context-based models of episodic memory.

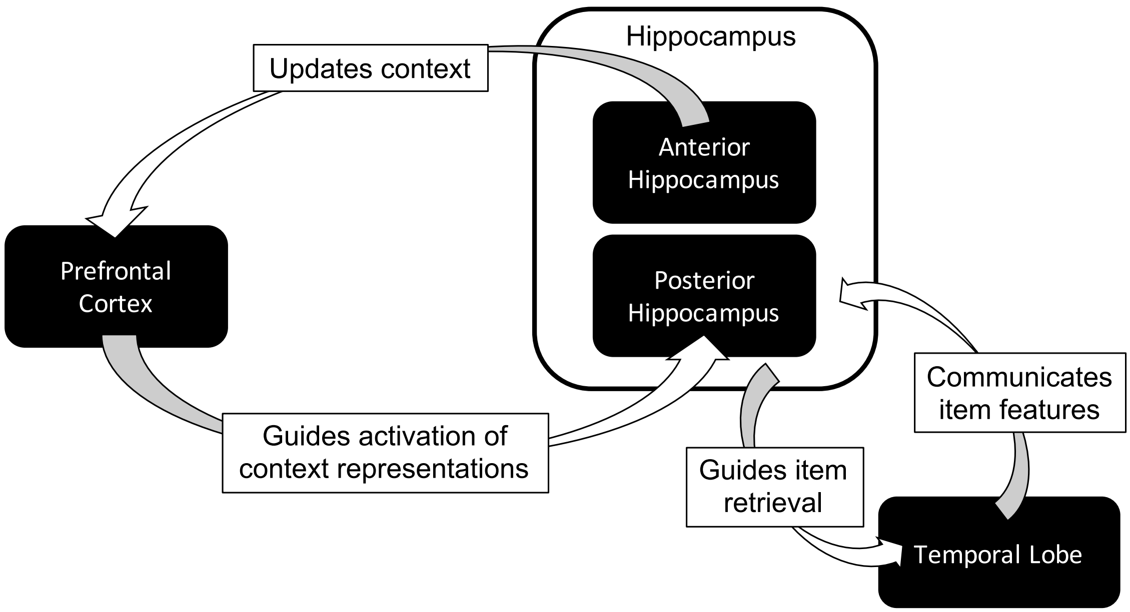

Figure 1.4 Visual representation combining the frameworks proposed by Polyn and Kahana (2008) and Eichenbaum (2017a) to explain the neural basis of cognitive control of episodic memory. The prefrontal cortex (PFC) is responsible for maintaining and updating the current state of temporal context, which includes a representation of task goals. The primary role of the hippocampus is to bind items to the current state of context and use those context-item associations to later retrieve items, whose identities are stored elsewhere in the temporal lobe. During encoding, the hippocampus communicates the current state of context to the PFC, allowing it to develop contextual rules and determine the task goals. During retrieval, the anterior hippocampus communicates the current state of context to the PFC, allowing it to activate the appropriate set of task goals. Then, the PFC uses the task goals to bias the retrieval of context in the posterior hippocampus. The activated context is then used as a cue to retrieve items, which are represented elsewhere in the temporal lobe.

The PFC Represents Task Goals as Part of Context

Models of cognitive control emphasize the role of the PFC in maintaining active representations of task goals even in the face of interference (e.g., Hazy et al., 2006; Miller & Cohen, 2001; O'Reilly et al., 1999; Wagner, 2002). The PFC is thought to be specialized for the maintenance of internal context information, including a representation of the current task goals (Jenkins & Ranganath, 2010; Miller & Cohen, 2001; O'Reilly et al., 1999). Direct evidence for the representation of task goals in the PFC comes primarily from literature on cognitive control. Lesions to the PFC impair animals' ability to represent abstract rules, and in non-lesioned animals, the PFC is highly active when abstract rules must be maintained, particularly across a delay (Navawongse & Eichenbaum, 2013; Rich & Shapiro, 2009; White & Wise, 1999; Wise et al., 1996). In humans, some regions of the PFC are highly activated in response to task-specific information about rules or goals. For example, when participants are required to switch between two different tasks, each set of task rules is represented in a different region of the PFC (Yeung et al., 2006).

Task goals may be represented in retrieved context models through an additional dimension of context. One computational implementation of retrieved context models, the Context Maintenance and Retrieval model (CMR; Lohnas et al., 2015; Polyn et al., 2011) includes not only a representation of temporal context but also a representation of semantic associations between words. This allows the model to make predictions for both temporal and semantic organization as well as how semantic organization may interact with temporal context to enhance or reduce the TCE. Further, CMR considers *context* to be made up of two components: *temporal context*, which contains a record of the mental activation across time and changes when studied items activate their own pre-existing associations, and *source context*, which represents other dimensions along which items may be similar, such as the task goals active when the item was encoded (Polyn et al., 2009a). More recent versions of CMR allow other kinds of associations to be represented in source context. As a result, the model can explain clustering based on item attributes such as emotional valence or associated reward (Horwath et al., 2023; Talmi et al., 2019; for a different approach to modeling spatial contiguity, see Howard et al., 2005).

Assuming that task goals are represented as a part of context, the same mechanisms should be responsible for both organization based on temporal order and organization based on task goals. Participants should therefore be able to organize their recalls based on task goals in the same way as they organize their recalls by temporal order. Polyn et al. (2009a) tested this prediction in an experiment where participants either completed the same encoding task for the entire list or switched processing tasks halfway through the list (task-shift lists). They found that participants tended to group their recalls by encoding task as well as by time. Polyn et al. (2009a) also successfully modeled participants' tendency to cluster their recalls by encoding task by representing task goals in the source context with CMR.

More direct evidence that task goals are represented as a part of mental context comes from neuroimaging studies, where task-specific patterns of brain activity can be identified. Polyn et al. (2012) measured participants' brain activity through fMRI during a free recall task and identified distinct patterns of neural activity associated with each of two different encoding tasks participants performed. When an item was retrieved, the pattern associated with its encoding task was also reinstated, providing evidence that information about task goals is represented as a part of mental context (Polyn et al., 2012; see also Sederberg, Schulze-Bonhage, Madsen, Bromfield, McCarthy, et al., 2007).

Prominent theories of cognitive control suggest that task goals can be represented by the PFC as a part of mental context, and these claims are supported by both behavioral and neural evidence. However, representation of goals is not enough. How do those goals influence other brain regions to guide memory processing?

The PFC Uses Task Goals to Bias Processing

Many theories of cognitive control emphasize that the role of the PFC is to bias activity in other brain areas, thereby indirectly influencing behavior to align with current task goals (Buckner & Wheeler, 2001; Miller & Cohen, 2001; O'Reilly et al., 1999). The PFC may control memory processes at encoding and retrieval by biasing activity in the hippocampus (Eichenbaum, 2017a; McClelland et al., 1995; Moscovitch, 1992). Such communication is possible because the PFC is highly connected to the hippocampus and surrounding areas, which are responsible for retrieving context information (for reviews, see Eichenbaum, 2017b; Preston & Eichenbaum, 2013; Simons & Spiers, 2003). Communication between brain regions can be inferred when they display similar patterns of activation, and activity in the PFC and the hippocampus is often correlated during memory encoding or retrieval (for a review, see Eichenbaum, 2017b). Specifically, activity in some PFC cells is highly correlated with theta oscillations (4-12 Hz) in the hippocampus during memory activities (Jones & Wilson, 2005). In animals, such theta synchrony between the PFC and hippocampus promotes memory on tests that require the retrieval of item-context associations (Benchenane et al., 2010; Hallock et al., 2016; Hyman et al., 2010; Kim et al., 2010; Place et al., 2016). By observing the timecourse of this correlation, we can also see which region influences activity in the other. When context is used to cue retrieval, changes in theta activity are initiated by the hippocampus, followed by changes in PFC theta activity. However, the direction of communication is reversed when task rules are applied to guide retrieval, and theta synchrony is instead led by the PFC (Hallock et al., 2016; Place et al., 2016). Not only does PFC-hippocampal communication promote successful retrieval, but these findings also indicate that activity in one region can directly influence activity in the other.

The idea that cognitive control of episodic memory is based on communication between the PFC and hippocampus is central to a recent model proposed by Eichenbaum (2017a; see Figure 1.4 for a visual of this model combined with the framework proposed by Polyn and Kahana, 2008). In Eichenbaum's (2017a) bi-directional communication model, the hippocampus and PFC communicate with each other during both encoding and retrieval. During encoding, their model assumes that the anterior hippocampus processes events occurring in the same context and binds them together to create a representation of the current context. This information is then communicated to the PFC, which develops a representation of the contextual rules. When the context associated with an event is reinstated, it is represented in the anterior hippocampus, which again communicates the context information to the PFC. In response, the PFC engages the context-appropriate rules, biasing retrieval in the posterior hippocampus towards context-appropriate memories. This model is rooted in the cognitive control literature, but the principles are highly relevant for developing retrieved context models because it sets clear criteria for how control processes should be implemented. A complete model should include mechanisms that allow task goals to influence context representations during both encoding and retrieval as well as an explanation for how those goals are developed.

The PFC May Bias Processing During Encoding

Task goals during encoding may influence the kinds of information encoded, thereby affecting later recall organization and the size of the TCE. Although participants in nearly all conditions display temporal contiguity, indicating that temporal context is encoded automatically to some extent, the TCE is reduced when participants are less able to engage in temporal encoding strategies (e.g., Mundorf et al., 2021). Therefore, both automatic and strategic control processes may play a role in the degree of temporal order information encoded, influencing later recall organization.

During encoding, the PFC may engage cognitive control by creating an attentional template which directs attention primarily to task-relevant details (O'Reilly et al., 1999; Summerfield, 2006; Wagner, 2002). Assuming attention is a limited resource (e.g., Kahneman, 1973), focusing attention only on task-relevant features should improve later memory performance (Benjamin, 2007). Focusing on some features may entail encoding less of other features. For example, participants who are instructed to focus on semantic meaning during encoding may encode less temporal information than those who focus on temporal order during encoding (Long & Kahana, 2017). The extent to which differences in recall organization can be attributed to strategic control processes at encoding compared to those at retrieval is still an open question. For example, the reduced temporal contiguity observed in lists with a semantic structure (Healey & Uitvlugt, 2019; Polyn et al., 2011) could be due to participants strategically encoding more semantic associations and fewer temporal associations, leaving less temporal information available when participants later search their memories and reducing the TCE. Alternatively, the same amount of temporal information may be encoded regardless of participants' encoding strategies, and the differences in later recall organization may be primarily due to strategic control processes operating at retrieval.

Although both of these interpretations are possible, there is at least clear evidence that the PFC is involved in encoding. Both the PFC and hippocampus are highly active during encoding, and greater activation in these regions during encoding predicts later retrieval success (Alkire et al., 1998; Blumenfeld & Ranganath, 2007; Brewer et al., 1998; Polyn & Kahana, 2008; Sederberg, Schulze-Bonhage, Madsen, Bromfield, McCarthy, et al., 2007; Staresina & Davachi, 2006; Strange et al., 2002; Wagner, 2002). This is especially true for tests that require retrieval of temporal context information (Jenkins & Ranganath, 2010). Activity in certain regions of the PFC during encoding also predicts memory for item-context associations compared to item memory alone (Dobbins et al., 2002; Summerfield et al., 2006), consistent with the claim that the PFC plays a role in encoding temporal context information. Given this evidence that the PFC is involved in encoding of episodic memories, a complete model of episodic memory must include mechanisms that allow for cognitive control during encoding.

The PFC May Bias Processing During Retrieval

A combination of automatic and controlled mechanisms also appear to be involved in the retrieval of temporal context. Both the hippocampus and PFC are highly active during retrieval, particularly when the memory test requires the retrieval of context information (Bonnici et al., 2012; Buckner & Wheeler, 2001; Eichenbaum, 2017b; Szczepanski & Knight, 2014). For example, both the hippocampus and anterior PFC are more highly activated during associative recognition, where participants are required to retrieve the association between two pair members, compared to an item memory task of similar difficulty (Ranganath et al., 2004). There is also evidence from the retrieved context models literature that the PFC is involved in context reinstatement. According to retrieved context models, when an item is retrieved it automatically reinstates its associated context from encoding, which then serves as a cue for the next retrieval. This context reinstatement has been observed in the PFC. Sederberg, Schulze-Bonhage, Madsen, Bromfield, Litt, et al. (2007) found that activity in the PFC reflected current task goals during encoding, and patterns of gamma activity in the PFC that predicted successful encoding were re-activated at retrieval.

In the cognitive control literature, extensive research has compared performance on memory tests for individuals with impaired cognitive control versus healthy controls. Populations with impaired cognitive control have selective deficits on tests that require context reinstatement and memory for temporal order. For example, older adults, who have diminished cognitive control and PFC volume relative to younger adults (Hasher et al., 2007; Raz et al., 1997; West, 1996), have greater deficits to recall than recognition and display reduced temporal contiguity in their recalls compared to younger adults (Hasher & Zacks, 1988; Howard et al., 2006; Kahana et al., 2002). Patients with lesions to their frontal lobe are also particularly impaired on tests of free recall, with fewer impairments on cued recall and recognition tests (Dimitrov et al., 1999; Moscovitch & Winocur, 1995; Stuss et al., 1994; Wheeler et al., 1995). If control processes are important for context reinstatement, then impaired cognitive control should also lead to impaired memory for context itself. Indeed, both older adults and frontal lobe patients display greater deficits in performance on explicit tests of context information than memory for individual items (Buckner & Wheeler, 2001; Butters et al., 1994; Duarte et al., 2005; Janowsky et al., 1989; Kuhl & Wagner, 2009; Milner et al., 1985; Naveh-Benjamin, 2000). Impaired cognitive control abilities therefore appear directly related to deficits in memory for context.

Integrating models of cognitive control with retrieved context models leads to another prediction: impaired PFC function should impact recall organization. Based on theories of cognitive control, the PFC is important for biasing retrieval towards task-relevant items. And under retrieved context models, recall organization is directly a result of the context that is reinstated when an item is retrieved. Therefore, if the PFC influences context reinstatement in the hippocampus, it should also influence recall organization. Consistent with this prediction, impaired cognitive control is associated with more disorganized recall. Deficits in temporal and semantic clustering are observed in frontal lobe patients (Gershberg & Shimamura, 1995; Hirst & Volpe, 1988; Incisa della Rocchetta, 1986; Jetter et al., 1986; Stuss et al., 1994; but see Alexander et al., 2003; Stuss et al., 1994) and in older adults compared to healthy younger adults (Healey & Kahana, 2016; Howard et al., 2006; Kahana et al., 2002). Interestingly, older adults may be able to compensate for baseline deficits in cognitive control by increasing activation of the PFC bilaterally (Cabeza et al., 2002; Talamonti et al., 2020). Declines in recall organization also predict cognitive decline beyond typical healthy aging. Individuals with a diagnosis of mild cognitive impairment (Malek-Ahmadi et al., 2011) or Alzheimer's (Gaines et al., 2006) exhibit reduced semantic clustering, and lower levels of both temporal and semantic organization predict later cognitive decline (Grober et al., 2000; Talamonti et al., 2021). These patients are less able to engage in cognitive control, which may contribute to their more disorganized recall.

Although most work focuses on the deficits in recall organization associated with frontal lobe damage, organization is not completely dependent on the frontal lobe and cognitive control. Mangels (1997), for example, found that temporal clustering was reduced for frontal lobe patients but not completely eliminated, leading him to conclude that temporal organization was influenced by both controlled and automatic processes. This is important from the perspective of retrieved context models, which predict that temporal information should be encoded and influence recall regardless of intentionality. Remembering an item should trigger the automatic retrieval of its associated context, which is inherently a better cue for other items experienced nearby in time. In Mangels's (1997) study, temporal organization was adaptive—using temporal order information at retrieval was helpful because few other memory cues were available. However, when temporal order information is not as useful, participants may adopt alternate strategies at retrieval and intentionally suppress temporal information, further reducing or even eliminating the TCE (Healey & Uitvlugt, 2019; Hong et al., 2022). Even if temporal order information was originally encoded, participants are able to strategically use other kinds of associations, like semantic associations, to guide recall instead. Such variations in retrieval strategy can be implemented in retrieved context models through differential weighting of associations. For example, CMR can account for both semantic and temporal contiguity in a list with semantic structure by representing semantic and temporal associations separately and including a parameter that biases retrieval towards relying on one associative dimension over the other (Morton & Polyn, 2016; Polyn et al., 2009a).

Cognitive Control Mechanisms in Models of Memory Organization

There is clear evidence that the PFC is important in developing and implementing retrieval strategies, and some work has endeavoured to represent these mechanisms in a computational model that has the capacity to make predictions about recall organization. For example, outside of the retrieved context models literature, some memory models include mechanisms for specific strategies, such as rehearsal during encoding or prioritizing some kinds of associations to guide retrieval (e.g., Lehman & Malmberg, 2013; Raaijmakers & Shiffrin, 1980, 1981). However, it is unclear if these models are able to account for the TCE observed when such control mechanisms are unlikely (such as during incidental encoding). Retrieved context models have been adapted to include some mechanisms for control, such as representing task goals as a part of mental context and post-retrieval monitoring (e.g., CMR; Lohnas et al., 2015; Polyn et al., 2009a). Recent work has also suggested mechanisms for retrieved context models through which participants can target retrieval from a specific context by reinstating the target context at the start of the retrieval period (Healey & Wahlheim, 2023; Zhang et al., 2022). Yet, the development of strategic control processes and their implementation to guide processing during encoding and retrieval are still underdeveloped in these models. Developing such mechanisms is important for explaining the effect of control processes on the TCE and other patterns of recall organization.

At least one model outside of the retrieved context models framework has used PFC-specific control processes to explain *semantic organization*. Becker and Lim (2003) proposed a model of cognitive control in which memories were represented in both a medial temporal lobe module, where item-context associations are stored directly, and a PFC module, where representations are weighted based on feedback provided after each recall to maximize the number of successful retrievals and minimize the number of repeats and intrusions. During retrieval, the representation of a recalled item is activated in the PFC module, which then cues the item-context association stored in the medial temporal lobe module. This activated item-context association is used to cue the next recall. Their PFC module can bias *which* associations are activated in the medial temporal lobe module by weighting its own representations to maximize performance. Becker and Lim (2003) demonstrated their model could produce significant semantic clustering in lists with semantic structure, fitting well to the behavior of healthy control participants. Lesioning the PFC module reduced both recall performance and semantic clustering, consistent with the behavior of actual patients with frontal lobe lesions. Because the purpose of this model was to simulate semantic organization, however, Becker and Lim (2003) did not test the model's ability to learn any other kind of recall organization. It remains to be seen how temporal recall strategies might be represented and applied in a model with both PFC and hippocampal components and if similar mechanisms can explain recall organization along other dimensions.

Overview of Experiments

A comprehensive theory of memory should include mechanisms to account for both automatic and controlled processes. However, to develop such a theory, we must distinguish which memory effects are automatic and which are controlled. The TCE is a good tool for making this distinction because there is clear evidence that temporal information is automatically encoded, and perhaps automatically retrieved, yet the size of the TCE is also modulated by strategic control processes. While retrieved context models provide a well-specified account of the automatic mechanisms underlying the TCE, they lack explicit mechanisms for control processes. Neural models of cognitive control provide guidance on how strategic control processes might be implemented in these models: representing task goals as a part of context and allowing task goals to bias context representations at encoding and retrieval.

However, it remains to be seen if and how control processes can be successfully implemented in retrieved context models. Additional empirical work is also needed to clarify the kinds of associative information that are automatically encoded and retrieved and the degree to which different strategies at encoding and retrieval affect recall organization. A better understanding of the behavioral effects of automatic and controlled processes on recall organization is important before these processes can be implemented in a computational model.

The experiments described in the following four chapters,[1] address several open questions: 1) Is temporal information automatically retrieved?, 2) How does assigning different tasks during encoding affect recall organization?, 3) To what extent do strategic control processes during encoding determine the availability of information for later recall organization?, and 4) Is the TCE present even in the presence of external, non-temporal goals?

Experiment 1 addresses the question of whether temporal information is automatically retrieved as well as automatically encoded by testing for a TCE in an implicit memory test, where retrieval is entirely automatic. The subsequent chapters turn to the effects of different kinds of control processes at encoding and retrieval on temporal contiguity, as well as other kinds of recall organization. Experiment 2 addresses the question of how different kinds of additional tasks that encourage different levels of processing during encoding might affect the TCE. Experiment 3 considers to what degree control processes during encoding determine the kinds of associations that are later available to guide recall by independently manipulating participants' encoding and retrieval strategies. Finally, Experiment 4 tested if the TCE occurs even in the presence of external, realistic goals beyond simply recalling as many words as possible by examining recall of items on a grocery list where each item was associated with a specific store where it could be purchased and a dish it would be used to make. Computational models were applied to the data of Experiments 3 and 4 to further test the predictions of retrieved context models and begin evaluating ways in which task goals could be presented in these models.

[1]Chapter 2 is adapted from "Incidentally encoded temporal associations produce priming in implicit memory" published in Psychonomic Bulletin & Review and is reproduced with permission from Springer Nature. Mundorf, A. M. D., Uitvlugt, M. G., & Healey, M. K. (2023). Incidentally encoded temporal associations produce priming in implicit memory. *Psychonomic Bulletin & Review*. https://doi.org/https://doi.org/10.3758/s13423-023-02351-w

Chapter 3 is adapted from "Does depth of processing affect temporal contiguity?" published in Psychonomic Bulletin & Review and is reproduced with permission from Springer Nature. Mundorf, A. M. D., Uitvlugt, M. G., & Healey, M. K. (2022). Does depth of processing affect temporal contiguity? *Psychonomic Bulletin & Review*, *29*, 2229–2239. https://doi.org/10.3758/s13423-022-02112-1

CHAPTER 2

IS TEMPORAL INFORMATION AUTOMATICALLY RETRIEVED IN AN IMPLICIT MEMORY TEST?

Experiment 1

A core assumption of retrieved context models is that the same automatic mechanisms that underly memory formation and retrieval naturally produce the TCE. According to these models, temporal information is automatically encoded whenever new memories are formed and automatically retrieved at recall (see Howard & Kahana, 2002; Howard et al., 2015; Polyn et al., 2009a). These models therefore can easily be extended to predict a TCE even under conditions of implicit retrieval, where memories are retrieved in the absence of any intent to retrieve or even the awareness that anything is being remembered.

Implicit memory is measured indirectly by examining the effect of previous experience on responses; for example, implicit memory can be inferred when responses to a repeated item are faster on its second presentation compared to its first presentation (repetition priming; Graf & Schacter, 1985). Repetition priming may be enhanced when a repeated target is preceded by a cue previously experienced nearby in the list (associative repetition priming). When words are studied in cue-target pairs, responses to a repeated target tend to be faster if it was preceded by its associated cue than if it was preceded by an unrelated item, even if the cue and target are not semantically related (McKoon & Ratcliff, 1979, 1986; Spieler & Balota, 1996). Some have suggested this associative repetition priming occurs because the cue and target form new associations during their first presentation; those associations are re-activated when the cue is repeated, facilitating responses to the target (Ratcliff & McKoon, 1988; Zeelenberg et al., 2003).

However, there are mixed findings on when, and if, associative repetition priming occurs (for a review, see Zeelenberg et al., 2003). Associative repetition priming is greatest when the cue and target are presented in the same order across presentations and at longer delays between repetition of the cue and repetition of the target (Raaijmakers, 2005; Zeelenberg et al., 2003). Associative repetition priming may even be eliminated if the presentation order is reversed, leading some to suggest this priming is due not to reinstatement of newly formed temporal associations, but rather to perceptual priming (Goshen-Gottstein & Moscovitch, 1995; Poldrack & Cohen, 1997) or unitization (encoding the pair as a single item; Graf & Schacter, 1989). Others attribute associative repetition priming to intentional retrieval strategies since the effect is largest when participants have more time between the repetition of the cue and repetition of the target (Carroll & Kirsner, 1982; Dew et al., 2007; Durgunoğlu & Neely, 1987; but see McKoon & Ratcliff, 1986).

One challenge in evaluating these accounts is that associative repetition priming has been investigated primarily in a paired associates paradigm. Little work has examined repetition priming among items not explicitly paired together (but see Smith et al., 1989). Therefore, it is not obvious if associative repetition priming occurs between *any* items merely experienced nearby in time or only between items explicitly paired together, and it is unclear if these associations are automatically retrieved.

Theoretical Predictions

Insofar as its mechanisms operate automatically, retrieved context models make clear, testable predictions for how temporal information should influence not only recall but also repetition priming. For recall, retrieved context models clearly predict temporal contiguity regardless of encoding intentionality (Mundorf et al., 2021) because they assume that at encoding, items automatically form reciprocal associations with the current state of a drifting mental context. Context changes during encoding as each item is studied while still retaining a record of the recent past, such that items studied relatively closer in time form associations with more similar states of context. When an item is retrieved, it reinstates its associated context from encoding. Because items studied closer together in time are associated with more similar states of context, the reinstated context tends to be a better cue for closer, relative to farther, temporal associates. In this way, temporal associates indirectly cue one another. Thus, retrieved context models naturally predict a TCE.

If we assume that the same context reinstatement occurs in implicit retrieval, these models clearly predict associative repetition priming for paired associates. Further, we can make two novel predictions regarding the conditions under which associative repetition priming should occur beyond a paired associates task. First, associative repetition priming should occur for items not explicitly paired together, even if they were originally separated by other list items, because the context associated with each item during encoding contains a record of the recent past. Second, because context changes with each item studied, the degree of priming should vary with initial *lag* (distance between cue and target on their first presentation). Current quantitative implementations of retrieved context models specifically predict more priming at shorter initial lags. In contrast, accounts which attribute associative repetition priming to perceptual priming, unitization, or intentional strategies predict associative repetition priming *only* for items explicitly paired together and always presented in the same order (initial *lag* = +1).

In the present experiment, participants read a series of words aloud, unaware that some would later be repeated or that their memory for the words would be tested. To avoid explicitly pairing words together, each word was presented individually at a regular interval. Only a small proportion of words were repeated, and the initial lag between repeated words was varied so participants could not predict which words would later be repeated together. This also allowed for a test of if the degree of associative repetition priming varied with initial lag.

Methods

Participants read 505 words aloud into a microphone as each word appeared one at a time on the screen. Most words were presented once, but the stimuli of interest were 30 pairs of words which were each presented twice (60 unique repeated words total). Each pair was composed of a cue word and a target word. After completing the reading task, participants were given 3 minutes for a surprise free recall test on all of the words they had read. The task took approximately 22 minutes to complete.

Participants

Because of the novel design of the current experiment, setting sample size through a precise *a priori* power calculation was not possible. However, as a general guideline I used Healey's (2018) finding that achieving 95% power to detect a TCE in *deliberate* memory search following incidental encoding requires a sample size of 510 participants per condition. Thus, a goal was set of collecting data from at least 500 participants. The final sample included 723 Michigan State University undergraduates who completed the experiment for course credit. Due to a technical error, demographic information was only recorded for 714 participants. Of these, 562 (78.7%) identified their gender as female, and the mean age was 19.7 years (SD = 1.9).

Data Exclusions

To eliminate the potential influence of intentional study strategies, data were excluded for participants who indicated on a post-task questionnaire that they suspected their memory would be tested. After making these exclusions, 603 participants (83.4%) remained.

Materials

Participants read 505 words, 385 of which were presented only once. The remaining 120 words were composed of 60 unique words each presented twice. These 60 words were divided into 30 pairs, where one member of each pair was designated as a cue word and the other as a target word.

The words were presented to participants as one long, continuous list. However, the experiment was in fact divided into 30 sections, or pseudo-lists, each composed of 10 to 25 words. Example pseudo-lists are presented in Figure 2.1. Each pseudo-list was composed of two words each presented twice (the cue and the target) and some number of once-presented filler words (represented with an X in Figure 2.1).

So that participants would not be able to anticipate, even unconsciously, when cue or target words would appear, each pseudo-list began with a jittered number of filler words (between 0 and 9). On their first presentation, I manipulated the initial lag between the cue and target—the distance in serial positions between the first presentation of the target word and the first presentation of the cue word ($lag = target - cue$). There were six possible initial lags: −5, −2, −1, +1, +2, or +5. Positive initial lags occurred when the cue was presented before the target; negative initial lags occurred when the target was presented before the cue. For example, if the cue was presented in serial position 1 and the target in serial position 2, the initial lag between them would be $2 - 1 = +1$ (Figure 2.1A). If instead the target was presented first in serial position 1, and the cue was presented in serial position 6, the initial lag between them would be $1 - 6 = -5$ (Figure 2.1B). Any serial

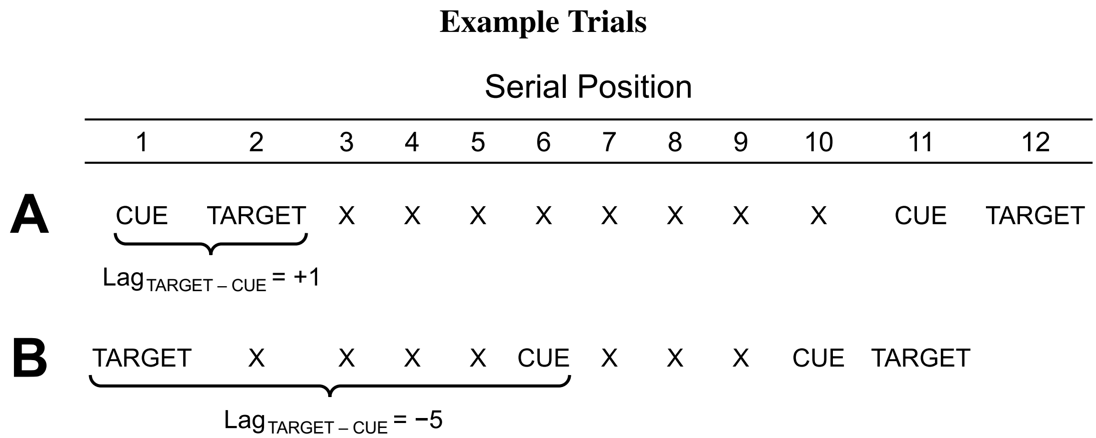

Figure 2.1 In these example trials, each X represents a once-presented filler word, and repeated words are labeled as either CUE or TARGET. The words were arranged into pseudo-lists, where each pseudo-list contained between 10 and 25 words. Each list began with a jittered number of filler words (between 0 and 9). There were six possible initial lags between the first presentation of the target and the first presentation of the cue: −5, −2, −1, +1, +2, or +5. On their second presentation, the cue was always presented first, immediately followed by the target. (A) An example pseudo-list. The cue is presented in serial position 1, immediately followed by the target in serial position 2. Here, the initial $lag = 2 - 1 = +1$. The inter-presentation lag between the first and second presentation of the cue = $11 - 1 = +10$. (B) An example pseudo-list where the initial $lag = 1 - 6 = -5$ and the inter-presentation lag for the cue = $10 - 6 = +4$. In all pseudo-lists, the inter-presentation lag for the target = +10.

positions between the first presentations of the cue and target were filled with once-presented filler items. In this example, the target is followed by four filler items (serial positions 2, 3, 4, and 5) before the first presentation of the cue. Each participant experienced each possible initial lag five times.

On their second presentation, the cue was always presented first, immediately followed by the target ($lag = +1$). Filler words intervened between repetitions such that there was always an inter-presentation $lag = +10$ between the first and second presentation of the target. Given that priming effects tend to be smaller at longer inter-presentation lags (Bentin & Moscovitch, 1988), the inter-presentation lag was held constant for the target words, which were the key stimuli of interest. This allowed for an examination of the effect of the initial *target − cue* lag on repetition priming for target items free from potential confounding effects of the target's inter-presentation lag. The inter-presentation lag for the cue word necessarily varied depending on the initial *target −*

cue lag.

For each participant, the 385 filler words were randomly drawn without replacement from a pool of 1,198 one- or two-syllable nouns containing between 2 and 9 letters, a subset of a larger word pool developed for the Penn Electrophysiology of Encoding and Retrieval Study (PEERS; Healey & Kahana, 2014; Siegel & Kahana, 2014). The repeated words were selected from a pool of 60 target words used by Healey et al. (2014) for a similar naming time task. These words were also one- or two-syllable nouns with between 2 and 9 letters. The cue and target words for each list were randomly selected without replacement from this pool for each participant. Thus, across participants, each word was equally likely to be chosen as a cue or target word.

Procedure

Participants completed the experiment individually in sound-insulated testing booths. Instructions appearing onscreen stated the experimenters were interested in developing a list of words for a future experiment that were neither too easy nor too difficult to process, and how quickly a person can initiate reading a word can be a measure of how difficult it is to process. Participants were therefore asked to read each word aloud *as soon as it appeared*. Participants were *not* informed that some words would be repeated or that they would be tested on the words at the end of the experiment. They were asked to avoid movement and making extraneous noise during the session (such as tapping their feet or coughing) because clear audio recordings were important for the experiment. Vocal responses were recorded for each word using a microphone placed in front of the computer. Participants were provided with two short breaks: one after the tenth pseudo-list and another after the twentieth pseudo-list. During these breaks, they could make noise and move around the booth if desired. Aside from these breaks, participants experienced the task as one continuous list.

In the reading time task, each word was presented individually on the screen for 1.5 s followed by a 500 ms inter-stimulus interval. Therefore, the stimulus onset asynchrony between any two words was 2,000 ms. After reading all 505 words, participants were given 3 min for a surprise free recall task. They were asked to type any words they could remember in whatever order they came to mind. Recalls were typed into an onscreen text box and submitted by pressing ENTER after each word. After pressing ENTER, the word they had just typed disappeared, leaving a blank text box for participants to type their next recall. Finally, participants were asked “At any point while reading the words, did you suspect you would be asked to remember the words later?” Participants were then debriefed and asked not to share the details of the incidental memory test with anyone else.

Data Scoring

Naming Time Detection and Reliability

To measure implicit memory, naming time (i.e., how long it took participants to begin reading the word aloud once it appeared onscreen) was compared for the first versus the second presentation of each repeated word. Chronset, a tool designed to distinguish noise from speech by analyzing recordings on the basis of multiple acoustic features, was used to detect speech initiation for each word (Roux et al., 2017). Chronset has been shown to be successful in detecting speech onset at rates similar to those of human raters (Roux et al., 2017). Both the first and second presentations were scored for all cue and target words using Chronset and MATLAB 2018. The same software was also used to score naming times for filler words for an additional followup analysis.

To verify that Chronset was accurately classifying speech onset, four human raters manually marked naming times for the first 41 participants. Each rater determined naming time for the second presentation of each target word using Audacity® recording and editing software (Version 2.3.0; Audacity Team, 2018). For any word where it was difficult to determine naming time (for example, due to poor recording quality), the human raters marked their ratings as low-confidence, and these low-confidence ratings were excluded from the reliability analysis. Only 1.5% of ratings were low-confidence.

Intraclass correlation (ICC) estimates and their 95% confidence intervals using a two-way mixed effects model to calculate variability among human raters are reported in Table 2.1. An ICC value of 0 indicates no agreement among raters, while an ICC value of 1 indicates perfect agreement (Shrout & Fleiss, 1979). Reliability was high among the human raters (ICC = 0.956 [0.952, 0.960]).

Chronset ratings were highly correlated with each of the human raters (Range: 0.872–0.903).

Naming Time Exclusions

Differences in response times tend to be quite sensitive to any fast or slow outlying data points (Ratcliff, 1979). Given that the measure of implicit memory relies on differences in naming times, I utilized exclusion criteria similar to that adopted by Healey et al. (2014) to remove outlying naming times. First, any responses faster than 200 ms or slower than 2,000 ms were eliminated. This step excluded responses on 0.6% of trials. After excluding these extreme outliers, additional exclusions were made for each participant based on their distribution of reaction times. After excluding these extreme outliers, values for each subject that were more than 2.5 standard deviations from the mean for each subject were also excluded. Means and standard deviations were calculated separately for each subject for each of the four response types (first presentation of the cue, first presentation of the target, second presentation of the cue, second presentation of the target). In total, 8.9% of all trials were affected, with no more than 20% of trials excluded for any single subject.[1]

Correlations for Human Raters and Chronset Ratings
Table 2.1 Correlations reported here are Pearson's r. All correlations were significant, $p < .001$.

	Rater 1	Rater 2	Rater 3	Rater 4	Chronset
Rater 1	–				
Rater 2	0.972	–			
Rater 3	0.983	0.965	–		
Rater 4	0.933	0.941	0.927	–	
Chronset	0.903	0.877	0.882	0.872	–

Results

Measures of Implicit Memory

I examined repetition priming for both cues and targets to test the predictions that 1) associative repetition priming should occur even when items are not explicitly paired, and 2) the degree of associative repetition priming should be affected by the initial lag between the cue and target.

[1] This approach to exclusions did not affect either the direction or significance of any analyses compared to making no exclusions at all.

Average Naming Times for Cue and Target Items
Table 2.2 All naming times are in milliseconds. For each item type, the mean naming time for each participant was calculated first, and then the mean and standard deviation (SD) were calculated across participants.

Item Type	Mean Naming Time (SD)	
	First Presentation	Second Presentation
Cue	519.11(63.02)	510.75(63.66)
Target	519.37(64.17)	498.34(61.90)

Average naming times are presented in Table 2.2. Repetition priming was measured for each repeated item by subtracting naming time on the first presentation from naming time on the second presentation, with negative values indicating repetition priming.

As expected, basic repetition priming occurred for both cue and target items (Figure 2.2A). Importantly, however, the size of this priming effect was influenced by item type. Repetition priming was greater for targets than cues, $t(602) = 9.91$, $p < .001$, $d = 0.404$, indicating significant *associative* repetition priming. Consistent with retrieved context models, cuing the target with another word previously presented nearby in time enhanced repetition priming even when the words were not explicitly paired.

One limitation of this design was that because the second presentation was always later in the list (exactly 10 serial positions for targets, 10 serial positions on average for cues), differences in response times could be influenced by practice or fatigue effects. To verify these differences in inter-presentation lags were not responsible for the associative repetition priming effect, the analysis of associative repetition priming was conducted a second time, restricting the analysis of cues to only those pseudo-lists with a cue inter-presentation $lag = +10$. Figure 2.2 displays the original analysis (including all pseudo-lists) alongside this restricted analysis. Repetition priming for targets was unchanged because the target inter-presentation lag was always +10. Mean repetition priming for cues was largely unchanged when only pseudo-lists with cue inter-presentation $lag = +10$ were considered although only 1/6 of trials were included for cues in the re-analysis. Differences in

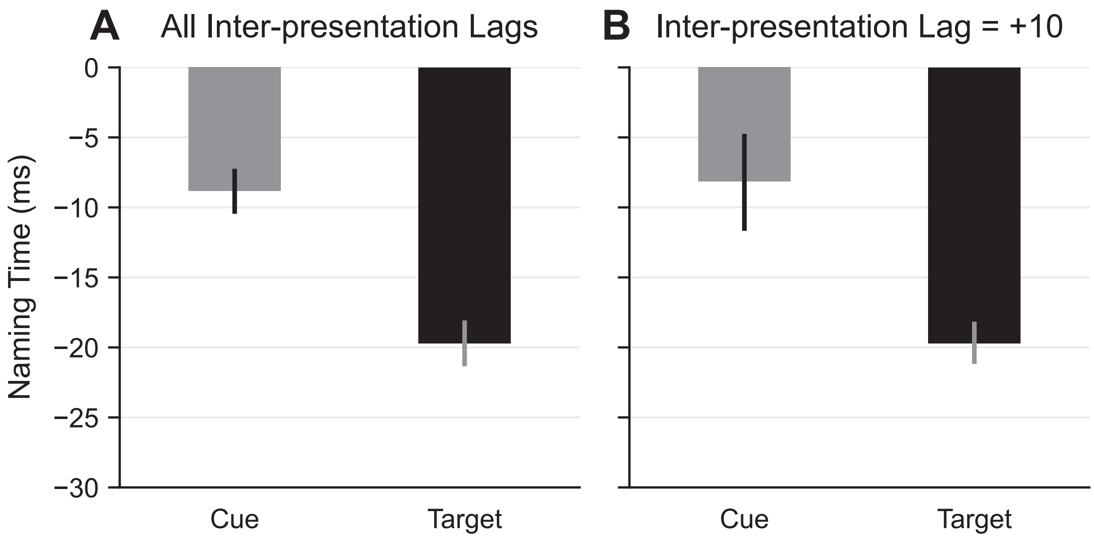

Figure 2.2 Average repetition priming for cue and target items. Repetition priming was calculated for each participant for each item as naming time on the item's second presentation minus naming time on the item's first presentation. (A) Average repetition priming from all pseudo-lists in milliseconds (ms), regardless of inter-presentation lags. (B) Average repetition priming from pseudo-lists with inter-presentation *lag* = +10 for each item type. Targets were always presented at inter-presentation *lag* = +10, so all pseudo-lists are included for targets. Cues were presented at inter-presentation *lag* = +10 on only 1/6 of pseudo-lists; only those pseudo-lists are included here for cues. Negative values indicate a repetition priming effect, and more negative values indicate a larger repetition priming effect. Error bars are bootstrapped 95% confidence intervals.

inter-presentation lag consistency for cues and targets did not affect repetition priming.

Temporal Contiguity Effects on Priming

To examine the effect of lag on associative repetition priming, priming of targets was examined for each possible initial lag (Figure 2.3). Repetition priming occurred at all initial lags, even when the cue and target were initially separated by other items (initial $|lag| = 2$ or 5) or their presentation order changed from the first to the second presentation (negative initial lags).

There was also an effect of initial lag (−5, −2, −1, +1, +2, or +5) on repetition priming using the Greenhouse-Geisser correction to account for a violation of the sphericity assumption, $F(4.86, 2921.90) = 4.05$, $p < .001$, $\eta^2 = .005$.[2] This difference was driven primarily by a *smaller* repetition priming effect at *lag* = +1, when the cue and target appeared in the same order on

[2]The by-lag analyses are based on the 602 participants who contributed data for all 6 possible lags.

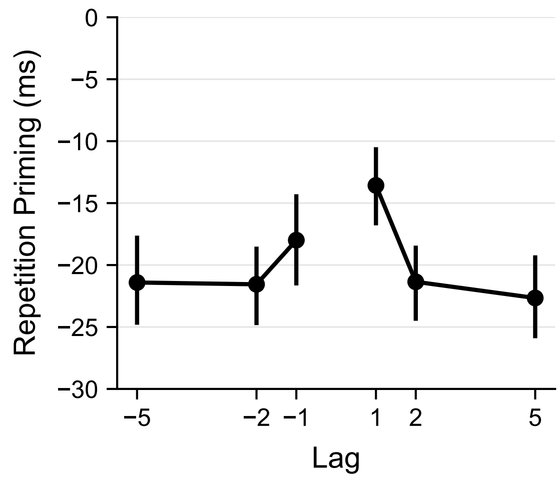

Figure 2.3 Repetition priming of target words plotted by the initial target minus cue lag in milliseconds (ms). Repetition priming was calculated for each participant for each item as naming time on the item's second presentation minus naming time on the item's first presentation. Negative values indicate a repetition priming effect, and more negative values indicate a larger repetition priming effect. Error bars are bootstrapped 95% confidence intervals.

both presentations (Figure 2.3). Planned contrasts revealed that repetition priming was reduced at $lag = +1$ relative to all other lags, $t(601) = -4.25$, $p < .001$. Most current implementations of retrieved context models assume that retrieving one item facilitates retrieval of other items studied closer, relative to farther, in time. If the same mechanisms are responsible for repetition priming we would expect greater, not reduced, repetition priming at $lag = +1$. Potential explanations for the lag effects are considered in the Interim Discussion.

Measures of Explicit Memory

Although my main focus was to test the predictions of retrieved context models for associative repetition priming, the free recall test is also of interest. If similar mechanisms underlie explicit and implicit retrieval, we would expect similar patterns to emerge in both memory tests. Analyzing temporal contiguity in recall also serves as an additional test of retrieved context models, which predict a TCE in almost any circumstance (see Healey et al., 2019). Since the TCE tends to be smaller for longer lists (Healey et al., 2019; Hong, Fazio, & Polyn, 2019), the small TCE previously observed in incidental encoding (Mundorf et al., 2021) may disappear altogether in a list of 505

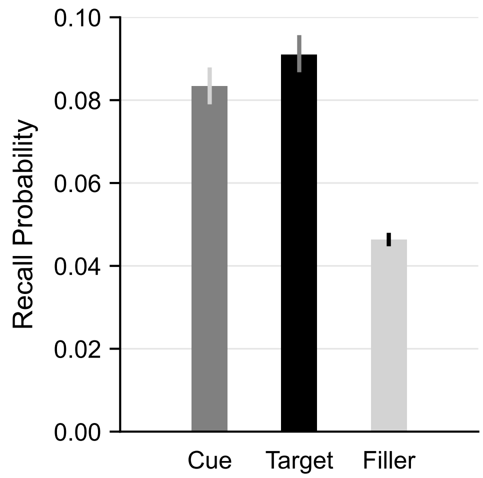

Figure 2.4 Recall probability for cue, target, and once-presented filler items on the final free recall test. For cue and target items, recall probability was calculated by dividing the number of cue or target items recalled by the number of unique cue or target items viewed during the naming task (30). For filler items, recall probability was calculated for each participant by dividing the number of filler items recalled by the total number of filler items viewed during the naming task (385). Error bars are bootstrapped 95% confidence intervals.

items.

Repetition Effects on Recall

Recall probabilities for the three item types (cue, target, and filler) are displayed in Figure 2.4. There was a significant effect of item type using the Greenhouse-Gesisser correction to account for a violation of the sphericity assumption, $F(1.67, 1002.78) = 246.36$, $p < .001$, $\eta^2 = .182$. Recall was higher for both cues, $t(602) = 20.17$, $p < .001$, $d = 0.821$, and targets, $t(602) = 23.00$, $p < .001$, $d = 0.937$, relative to once-presented filler items with a Bonferroni-adjusted $\alpha = .05/3 = 0.017$. Although no strong conclusions are based on these contrasts because the repeated and filler items were selected from a different (albeit similar) word pools, these results are consistent with previous findings that repeated items are better remembered (e.g., Glanzer, 1969). Target items were also slightly more likely to be recalled than cue items, $t(602) = 2.56$, $p = .011$, $d = 0.104$. In both recall and repetition priming, memory was better for target than cue items.

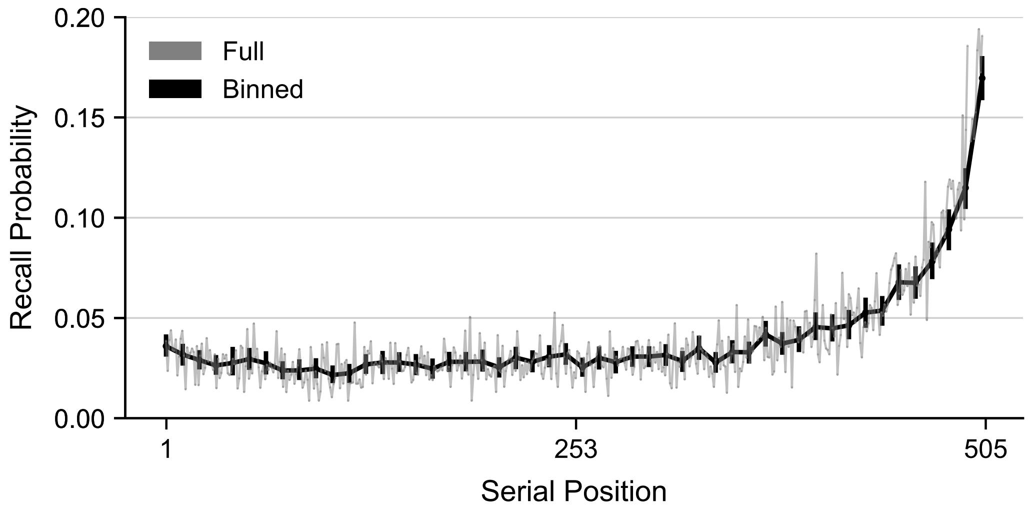

Figure 2.5 Full (gray) and binned (black) serial position curves for recall of once-presented filler items. These serial position curves (SPCs) represent the probability that a once-presented item from each serial position would be recalled, given that a once-presented item was actually presented in that serial position. The three serial positions which were only ever occupied by a repeated item (495, 504, and 505), were treated as missing values. Each point on the binned SPC represents the average recall probability for each bin of 10 serial positions except the last bin, which is an average of the final 15 serial positions. For example, the first point on the binned SPC represents the average recall probability for serial positions 1–10, and the second point represents the average recall probability for serial positions 11–20. Error bars are bootstrapped 95% confidence intervals.

Recall Dynamics

To facilitate comparisons between the current results and previous work examining free recall of only once-presented items, the detailed analyses of recall dynamics focus only on once-presented filler items. Serial position curves (SPCs) provide a measure of which items tend to be recalled. Figure 2.5 displays both the full SPC and a binned version. The full SPC plots recall probability for each serial position in the full list of 505 items. I also calculated a binned SPC to better visualize general trends such as primacy or recency. Each point on the binned SPC represents average recall probability for a bin of 10 consecutive serial positions (except the last bin, which is an average of the final 15 serial positions). Both SPCs display fairly low recall (see also Figure 2.4) with a strong recency effect, which is typical in immediate free recall of shorter lists (Glanzer & Cunitz, 1966; Ward et al., 2010).

Temporal Contiguity Effects on Recall

Temporal bias scores and temporal factor scores were used to measure temporal contiguity in recall. Temporal bias scores, introduced by Uitvlugt and Healey (2019), are similar to the lag-CRP, which gives the probability of making a transition of each lag conditional on the item at that lag being available (for details on how CRP is calculated, see Healey et al., 2019). Here, *lag* refers to the distance in serial positions between the just-recalled item and the next recall, not the distance between the cue and target on their initial presentation. For example, if a participant just recalled the item in the 3^{rd} serial position on the study list and then recalls the item from the 5^{th} serial position, that would be a transition of $lag = 5 - 3 = +2$. Temporal bias scores differ from the lag-CRP in that they can account for potential confounds from serial position effects by comparing the actual bias for making a transition of a given lag to a calculated chance level. Temporal bias for a given lag is calculated for each participant by counting the number of times a transition of that lag was actually made (actual count) and the number of times a transition of that lag would be expected to occur if items were recalled in random order (expected count; determined by permuting the order of recalls many times and counting the number of times a transition of that lag was made across permutations). The temporal bias score is simply $\frac{\text{actual count - expected count}}{\text{expected count}}$, calculated for each participant and each lag. Cases where both the actual and expected count were zero are treated as missing values. A score above zero for a given lag indicates it occurred more often than expected by chance, and a score below zero indicates it occurred less than expected.

Temporal factor (TF) scores are a single-number measure of the TCE that considers the *lag*, or distance, in serial positions between successively recalled items. TF scores are calculated for each list by taking the $|lag|$ of each transition made by a participant, finding its percentile within the distribution of all possible $|lags|$ for that transition, and then averaging across transitions (Polyn et al., 2009a; Sederberg et al., 2011). This analysis ignores the direction of the transition (forward or backward). Transitions outside the list boundaries or to previously recalled items are not considered possible. For example, $lag = +1$ would not be possible if the just-recalled item was the last item in the list. Higher temporal factor scores indicate near-lag transitions are more likely than far-lag

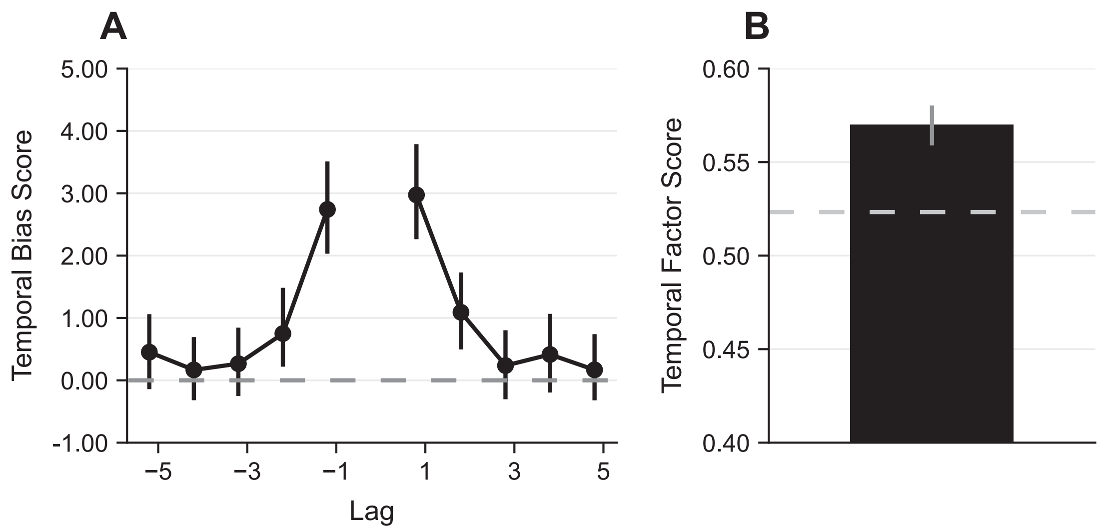

Figure 2.6 Temporal contiguity in recalls. (A) Temporal bias scores for recall of once-presented filler items, from *lag* = −5 to +5. Here, lag refers to the distance in serial positions between the just-recalled item and the next recall. Temporal bias scores for each lag were calculated by comparing the number of times a transition of that lag was actually made to the number of times it would be expected to occur by chance. Chance (expected count) was calculated by permuting the order of recalls for each list 500,000 times, counting the number of times a transition of each lag was made across permutations, and dividing by the number of permutations to get the number of times a transition of that lag would be expected to occur if the items were recalled in random order. For each participant and for each lag, the temporal bias score was calculated as (actual count − expected count) ÷ expected count. The dotted line indicates a score of zero (no bias). (B) Temporal factor scores for recall of once-presented filler items. For each transition, the percentile score of the actual $|lag|$ in a distribution of all possible $|lags|$ is calculated, where any transition to another item in the list that has not already been recalled is considered possible. Scores are calculated individually for each participant. Chance for temporal factor scores was determined by calculating temporal factor scores for each of 100,000 random permutations of the order of recalls and getting the average of this distribution. Chance temporal factor scores are represented with a dotted line. Error bars are bootstrapped 95% confidence intervals.

transitions (i.e., greater temporal contiguity). To control for primacy, recency, and other serial position effects, which may artificially inflate the TCE, the actual TF score was compared to the score expected if transitions were random with respect to lag. This chance-level expected factor score was calculated using the same permutation procedure described above.

As presented in Figure 2.6A, the temporal bias scores were greatest for near lags, particularly $|lag| = 1$. This replicates previous findings of a symmetrical TCE following incidental encoding, consistent with retrieved context models (Mundorf et al., 2021). TF scores allow for an even more straightforward test of the size of the TCE. The TCE was significantly greater than chance,[3] $t(598) = 10.75$, $p < .001$, indicating that participants did display significant temporal recall organization even after incidental encoding of a very long list.

Interim Discussion

The aim of this experiment was to test the predictions of retrieved context models with the assumption that temporal information is encoded and retrieved automatically. These models predict a TCE in free recall regardless of encoding intentionality; because temporal associations are encoded automatically, temporal contiguity should be observed even if participants are not intentionally studying. Consistent with this prediction, there was a TCE in recall following incidental encoding.

If we assume context reinstatement is automatic even when items are not intentionally *retrieved*, retrieved context models also make clear predictions for associative repetition priming. First, associative repetition priming should occur even when items are not presented as a pair, and second, this effect should vary with initial lag. These results are clearly consistent with the first prediction. Associative repetition priming occurred even when participants incidentally encoded the items without knowing which items would be repeated together later, indicating that new temporal associations were encoded automatically. Associative repetition priming also occurred among cue-target pairs originally separated by filler items or whose presentation order was reversed. This supports the models' assumption that during encoding, mental context retains a record of the recent past, so items separated by a few serial positions may still be associated with somewhat similar contexts and thus automatically cue one another at retrieval. Consistent with the second prediction, there was a small yet significant effect of initial lag on priming. This difference was primarily driven by reduced repetition priming when the cue and target were presented in the same order at both presentations (initial $lag = +1$).

Although these findings are generally consistent with retrieved context models, additional mechanisms not implemented in current recall-oriented models may contribute to this pattern of results. Current computational implementations of retrieved context models assume that items

[3] TF scores cannot be calculated for participants who did not make at least one valid transition, so those participants were excluded from this analysis.

experienced closer in time become associated with more similar states of context. A natural prediction is that a cue and target should be better primes for each other when they are initially presented in adjacent serial positions. However, the opposite pattern was present here. Why might this be?

Repetition priming may have been reduced at initial *lag* = +1 because participants consciously recognized words were being repeated. Given that conscious recollection may be more time-consuming than implicit retrieval (Dew & Cabeza, 2011; Jacoby, 1991; but see Dewhurst et al., 2006), responses may have been slower when the cue triggered conscious retrieval of the target. Participants may have also been surprised or confused upon detecting a repetition, further delaying their response. Conscious recollection was more likely on trials with an initial *lag* = +1 because the cue and target were repeated in exactly the same order, providing the strongest possible retrieval cue for the repeated target. Even if this occurred only on a few trials, average naming time for targets at *lag* = +1 would be slower. Implementations of retrieved context models assume that as an item's accessibility increases, responses to that item will be faster. It is possible that, instead, activation increases an item's accessibility up to a point (resulting in greater priming). Beyond that point, high activation may engage additional processes such as episodic or pre-episodic recollection (Smith et al., 2013), slowing down responses as observed here.

Another possibility is that targets were named more quickly on their first presentation if they followed an item from the cue/target word pool (initial *lag* = +1) versus following an item from the once-presented item word pool (all other initial lags). Because repetition priming was calculated as the naming time on the second – first presentation, faster naming of targets on their first presentation could lead to reduced repetition priming on initial *lag* = +1 trials. These are only two possible explanations; future work should test these and other possibilities to better understand the effect of initial lag on priming.

These results are inconsistent with several alternate explanations of associative repetition priming. If associative repetition priming is due to priming of items' perceptual features or unitization, repetition priming should be enhanced only if the cue and target were always presented in the same order (initial *lag* = +1). Similarly, if associative repetition priming is a result of intentional strategies, priming should be enhanced primarily for cue-target pairs consistently presented at *lag* = +1. The present design made it unlikely that participants could enhance repetition priming through encoding strategies, and, in fact, these results reveal the opposite pattern. Associative repetition priming occurred at all initial lags and was *reduced* at initial *lag* = +1.

This experiment supports retrieved context models' assumption that temporal information is both automatically encoded and automatically retrieved. A TCE was observed in free recall following incidental encoding, replicating previous findings that temporal information is automatically encoded. Associative repetition priming was observed even among items that were not explicitly paired together, indicating that temporal associations are also automatically retrieved. However, priming was *reduced* if the cue and target were presented close in time and in the same order on both presentations, contrary to specific models' predictions. The finding that new associations are automatically formed during encoding, allowing items studied nearby in time to later automatically cue one another, provides a strong foundation for examining the effects of strategic control processes on temporal organization. Given that temporal information is both automatically encoded and automatically retrieved, at least to some extent, what is the role of control processes in changing the size and shape of the TCE?

CHAPTER 3

DOES THE DEPTH OF A PROCESSING TASK DURING ENCODING AFFECT TEMPORAL CONTIGUITY?

Experiment 2

The previous chapter provides strong support for retrieved context models' assumption that temporal order information is both automatically encoded and automatically retrieved. However, even if order-based strategies are not entirely responsible for the TCE, the temporal organization observed in free recall may still be influenced by strategic control processes. As discussed in Chapter 1, greater temporal contiguity is typically associated with better memory performance (Healey et al., 2019; Sederberg et al., 2010; Spillers & Unsworth, 2011). This relationship between the TCE and recall is a natural prediction of retrieved context models; if the TCE and recall are a result of the same automatic mechanisms, then greater recall should be associated with greater temporal contiguity. However, the correlation between the TCE and recall is reduced when participants have less opportunity to engage in order-based encoding strategies (Healey & Uitvlugt, 2019; Long & Kahana, 2017; Mundorf et al., 2021), suggesting that control processes may also contribute to the TCE-recall correlation.

Another variable that is associated with better memory is depth of processing during encoding. Memory tends to be better for items processed according to meaning (deep processing) rather than perceptual features (shallow processing). This levels of processing (LOP) effect has been consistently observed in both recall and recognition regardless of encoding intentionality or specific deep processing task (Craik & Tulving, 1975; Eysenck, 1979; Hyde & Jenkins, 1969; Moscovitch & Craik, 1976; but see Rose & Craik, 2012). Extensive work has investigated interactions between deep processing and other aspects of memory, such as primacy and recency (Mazuryk & Lockhart, 1974) and semantic organization (Einstein & Hunt, 1980; Hyde & Jenkins, 1969). The benefits of deep processing have inspired recommendations for teaching methods, study strategies, and textbook design (Ayçiçegi-Dinn & Caldwell-Harris, 2009; Biggs, 1978; Martin et al., 1985; Seiver et al., 2019). Yet, the mechanisms through which deep processing influences memory are still not well understood (Baddeley, 1978; Craik, 2002; Eysenck, 1979).

Both LOP and the TCE have strongly influenced memory theory development, and both point to practical ways of improving memory. Yet, little work has examined how these effects interact. Theories which make the same predictions for summary measures, like overall recall, often make divergent predictions for the TCE, making temporal contiguity a useful tool for theory testing. Considering these two effects together allows us to develop a more unified theory of memory that can explain not only each effect independently but also how they interact. Below, I outline theoretically motivated hypotheses of how LOP might influence the TCE.

Reasons to Predict Deep Processing May Increase the TCE

Deeper LOP and a larger TCE are both associated with better recall. Thus, on purely empirical grounds, a reasonable hypothesis is that deeper processing should be associated with increased temporal contiguity.

Retrieved context models provide a theoretical motivation for this hypothesis. These models assume memories form when items become associated with the current state of a mental context representation which drifts through a high-dimensional representational space. When an item is studied, it activates its pre-existing associations, the activation of previous items' representations fade, and context drifts towards this just-studied item's representation. In this way, items studied nearby in a list become associated with similar states of context. When an item is recalled, it reinstates its associated context from encoding, providing a cue for items originally studied nearby in time. This naturally produces a TCE.

The size of the TCE depends on how far context drifts with each event. If items weakly activate their pre-existing associations, context will drift very little; all items will form associations with a similar state of context, and the TCE will be small. If each item strongly activates its pre-existing associations, mental context will drift farther toward the just-studied item's representation. Only items studied close in time will share similar contexts, resulting in a greater bias for transitioning between adjacent items which will enhance the TCE. In this light, deep processing should cause context to drift farther than shallow processing because a deep processing task involves not only activating items' perceptual features (as shallow processing does) but also deeper semantic features (as suggested by Healey & Kahana, 2016).

However, there is another possible interpretation of how LOP influence contextual dynamics. These models make a distinction between item and context representations. If deep processing acts primarily on item representations and not context, deeper processing would not increase the TCE. Examining the TCE under deep processing will help adjudicate between these competing interpretations of retrieved context models.

Reasons to Predict Deep Processing May Decrease the TCE

Other perspectives suggest deep processing should reduce the TCE. Under the item-order framework (Engelkamp & Zimmer, 1997; Hirshman & Bjork, 1988; Nairne et al., 1991), recall depends on processing information about individual items and inter-item associations like temporal order. But there is a trade-off: any manipulation that encourages item-specific processing should improve memory for specific items at the expense of memory for order. Thus, the TCE should be reduced (Lazarus et al., in prep; McDaniel & Bugg, 2008). For example, McDaniel et al. (2011) found a smaller TCE for lists of orthographically distinct items (e.g., *khaki, lynx*) compared to common items (e.g., *cookie, ruler*) and suggested the reduction was due to distinct words requiring more item-specific processing. Similarly, deeper processing may draw more attention to item-specific information (McDaniel & Bugg, 2008). The item-order account predicts deep processing should lead to better memory for items but reduced memory for order.

Finally, LOP may change participants' encoding strategies. Absent any experimenter-imposed encoding task, participants often adopt effective order-based strategies, such as linking items together to form a story (Delaney & Knowles, 2005; Hintzman, 2016; Unsworth, 2016). Such strategies may contribute to the TCE by encouraging serial recall (Bouffard et al., 2018; Unsworth et al., 2019). For participants using order-based strategies, *any* experimenter-imposed processing task that encourages focusing on individual items should interfere with such strategies, reducing recall and the TCE. That is, even if not all participants use order-based strategies, the average TCE should be highest with no encoding task. One study found deep processing reduced recall and the TCE relative to no task (Long & Kahana, 2017), but more work is needed to replicate these findings and compare both deep and shallow processing to no-task. The impact of encoding tasks on recall, on the other hand, likely depends on individual differences in the effectiveness of strategies employed. A task may not impair memory if participants are using ineffective strategies. Indeed, several studies report better recall for deep processing or no effect of task (Hunt et al., 2011; Hyde & Jenkins, 1969), while others report deep processing impairs memory relative to no task (Hagen et al., 1970; Mazuryk & Lockhart, 1974).

In sum, there are theoretically motivated reasons to suspect deep processing may increase or decrease the TCE. Existing literature lacks information on which hypothesis is accurate. Here, I propose to fill this gap.

Methods

The hypotheses, methods, and analysis plan for this study were preregistered prior to data collection (https://osf.io/4abjv/?view_only=f246b1d2f32d49f898f43e20fb045465; Healey et al., 2020).

In this experiment, participants studied 30 lists of words for free recall: 10 lists with no encoding task, 10 with a shallow encoding task (judging if the letter "T" was in the word), and 10 with a deep encoding task (judging if the word referred to a living thing).

Participants

Participants were Michigan State University undergraduate students who completed the experiment for course credit. Data collection began in September 2020 when Michigan State's classes were conducted remotely due to COVID-19. Therefore, all participants completed the study online.

Sample Size and Stopping Rule

A target sample size of at least 327 was selected to provide 95% $1-\beta$ power to detect a small effect ($d \geq 0.2$) via a two-tailed paired-sample t-test. Data collection was originally planned to end once the target sample size had been reached or at the end of the Fall 2020 semester, whichever came first. However, COVID-19 created a higher than normal demand within the department for online studies to allow students to meet course requirements remotely. To help meet this demand, data collection continued for the entire semester even after surpassing the original target. The data from existing participants were *not* examined prior to making this decision. In total, 825 participants completed the experiment.

Data Exclusion and Final Sample

Eight participants were excluded for not meeting the demographic exclusion criteria: three for reporting English was not their first language, four for failing to report their first language, and one for indicating they were over 18 at one point and under 18 at another point within the same session. For the remaining participants, data was excluded for any *list* where they recalled fewer than two list items (measuring the TCE requires at least two recalled items) or output more than 32 responses (i.e., twice the list length). Any *participant* who had more than 10% of their lists excluded (> 3 out of 30) was completely excluded from analysis. In total, 145 participants were excluded. This high exclusion rate reflects an overall low average performance in the sample. Among included participants, a total of 427 lists were excluded (71 from deep lists, 278 from shallow lists, and 78 from no-task lists).

The final sample included 680 participants (82.4% of the total sample); 470 were female, and the mean age was 19.6 (SD = 1.9). Participants in the final sample had an average of 97.9% of their lists included (SD = 3.1%, $Mode$ = 100%).

Materials

Participants studied 30 lists each composed of 16 words in an immediate free recall task. Lists were composed of words randomly selected from the pool of 1,638 nouns developed for PEERS (see Healey et al., 2019). Ten of the 30 lists were randomly assigned to each of the three conditions. Lists were presented in random order with the restriction that no more than two lists from the same condition were presented successively.

Before studying the first list, participants were given instructions explaining each encoding task and the free recall test that would follow each list. For each word in the the deep processing lists, participants were asked “Does this word refer to a living thing?” For the shallow processing lists, they were asked "Does this word contain the letter T?" Participants pressed the Y key for YES or the N key for NO while the word was on the screen. For the control no-task condition, participants were assigned no encoding task, were not required to make any keypress, and were free to study the words as they chose.

The letter "T" was chosen as the target letter for the shallow processing task in an effort to roughly match the expected number of YES responses in the deep processing task. To determine how many YES responses would be expected in the deep processing task, two undergraduate research assistants (i.e., from the same student body as the participants) and another lab member independently rated each of the words in the pool as either living or non-living. The three raters agreed for 1,425 out of 1,638 words. Some words were more difficult to judge than others; for example, the word *chest* might be judged as living if it is interpreted as a body part but judged as non-living if it is interpreted as a container (like a *treasure chest*). For the 213 words where they disagreed, two remaining lab members each made a YES/NO judgment and the modal judgment across all raters was taken as the expected response. For the deep processing task, 36.0% of the 1,638 words had an expected YES response. "T" occurs in 36.1% of the words in the pool, closer to 36.0% than any other letter.

Procedure

Each trial began with an instruction screen informing the participant which encoding task to perform for the upcoming list. To allow participants to take short breaks as needed, the instruction screen did not advance until the participant pressed SPACE. During the study phase, words were presented individually in the center of the screen for 1 s followed by a 400-600 ms jittered inter-stimulus interval. In deep and shallow lists, the relevant question was displayed above the word until participants entered a response. Then, the prompt disappeared, leaving just the to-be-studied word for the remainder of the 1 s presentation period. Following the presentation of the final word, participants had 60 s to recall as many words from the list as possible in whatever order they came to mind. Recall instructions were displayed onscreen throughout the recall period. Responses were typed individually, and participants were instructed to press ENTER after each response to submit it and clear the screen for the next response. Once the recall period had elapsed, instructions for the next list were presented.

Analyses

Temporal Contiguity

Temporal factor (TF) scores served as the primary measure of the TCE for this experiment because they provide a single-number measure of the TCE that can be compared to a calculated chance value. Comparing to a calculated chance value provides this analysis with resistance to confounding serial position effects such as primacy or recency, which may artificially inflate the TCE (described in detail in Chapter 2). To facilitate easier comparisons among conditions, in this experiment chance was subtracted from the actual TF scores. The difference was divided by the standard deviation of the chance distribution for each subject to provide a single number, referred to as *chance-adjusted TF scores*. Higher chance-adjusted TF scores indicate near-lag transitions are more likely than far-lag transitions (i.e., greater temporal contiguity), where *lag* is the distance between two words in their original serial positions in the list. This analysis ignores the direction of a transition (forward or backward).

Lag-conditional response probabilities (lag-CRPs) and temporal bias scores were also used to help visualize the TCE. Lag-CRPs give the probability of making a transition of each lag conditional on the item at that lag being available. Temporal bias scores are similar to the lag-CRP. However, they remove serious potential confounds from serial position effects in the same way as the chance-adjusted TF scores (for a detailed description, see Chapter 2). For this reason, I primarily rely on temporal bias and chance-adjusted TF scores as measures of the TCE. A temporal bias score above zero for a given lag indicates it occurred more often than expected by chance, and a score below zero indicates it occurred less than expected.

Semantic Contiguity

The analyses of temporal contiguity described above were part of a preregistered analysis plan. After conducting those analyses, I also conducted a set of followup analyses examining semantic contiguity, which is the tendency for words that are more strongly semantically related to be recalled together, to determine if LOP also affected semantic organization. Semantic relatedness between words was defined as the cosine of the angle between their high-dimensional vector representations in Word Association Space (WAS; Steyvers et al., 2004). Measuring word relatedness with WAS $cos(\theta)$ allows even small differences in word relatedness to be measured, which is important given that the present experiment used lists composed of randomly selected words. A measure analogous to chance-adjusted TF scores was used to quantify semantic contiguity. Chance-adjusted *semantic* factor (SF) scores are calculated in the same way as their temporal counterparts except that semantic lags are used instead of temporal lags. For a given transition, a semantic lag of 1 means transitioning to the most semantically similar available item in the list (in terms of $cos(\theta)$), a semantic lag of 2 means transitioning to the second most similar available item, and so on.

Results

Overall Recall

Probability of recall is displayed in Figure 3.1A. Mean recall was below 0.4 in every condition, lower than in past research with similar participants (e.g., Healey & Uitvlugt, 2019) but not unusual for intentional free recall using LOP instructions (e.g., Craik & Tulving, 1975; Hunt et al., 2011). Because the primary analyses involve relative differences among conditions, low recall should not impact interpretation of the results.

Planned pairwise tests revealed higher recall for no-task ($M = 0.368$, $SE = 0.005$) than either deep ($M = 0.316$, $SE = 0.003$), $t(679) = 15.10$, $p < .001$, $d = 0.579$, or shallow ($M = 0.271$, $SE = 0.003$) processing, $t(679) = 28.24$, $p < .001$, $d = 1.083$. This pattern is consistent with some past work where no-task participants displayed higher recall than either deep or shallow processing (Hagen et al., 1970; Long & Kahana, 2017; Mazuryk & Lockhart, 1974). There was also an LOP effect; recall was higher under deep than shallow processing, $t(679) = 24.43$, $p < .001$, $d = 0.937$.

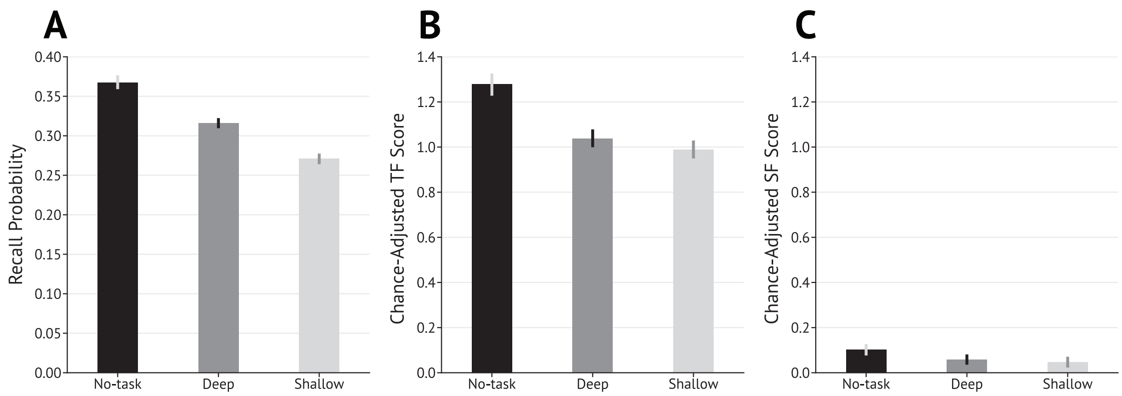

Figure 3.1 Measures of overall recall, temporal contiguity, and semantic contiguity for all conditions. (A) Recall probability, (B) chance-adjusted temporal factor (TF) scores, and (C) chance-adjusted semantic factor (SF) scores for no-task, deep processing, and shallow processing lists. For TF and SF scores, chance was determined by permuting the order of recalls 500 times. Scores were calculated for each list by subtracting the average of the chance distribution from the actual TF or SF score and then dividing by the standard deviation of the chance distribution. Error bars are bootstrapped 95% confidence intervals.

Temporal Contiguity

Chance-adjusted TF scores were above chance in all conditions (Figure 3.1B). Planned comparisons revealed a greater TCE for no-task (M = 1.28, SE = 0.03) than deep (M = 1.04, SE = 0.02), $t(679) = 10.93$, $p < .001$, $d = 0.419$, or shallow (M = 0.99, SE = 0.02) processing, $t(679) = 13.39$, $p < .001$, $d = 0.514$. The TCE was greater for deep than shallow processing, $t(679) = 2.87$, $p = .004$, $d = 0.110$, demonstrating an LOP effect on the TCE. This effect, though significant, was small. The size of this effect is worth noting, and I return to this issue in the Interim Discussion.

Recall Dynamics Curves

Although the main focus of this experiment is overall recall, the TCE, and their relationship, more detailed measures of recall dynamics may provide additional insight into how LOP influence memory search. Serial position curves measure recall as a function of serial position, and probability of first recall curves measure which serial positions are recalled first (Figure 3.2A and B). Recency was pronounced in all conditions, albeit larger for deep and shallow processing. Primacy was pronounced only for the no-task condition. This pattern is consistent with previous work where imposed processing tasks reduced primacy (e.g., Hagen et al., 1970; Long & Kahana, 2017; Mazuryk & Lockhart, 1974).

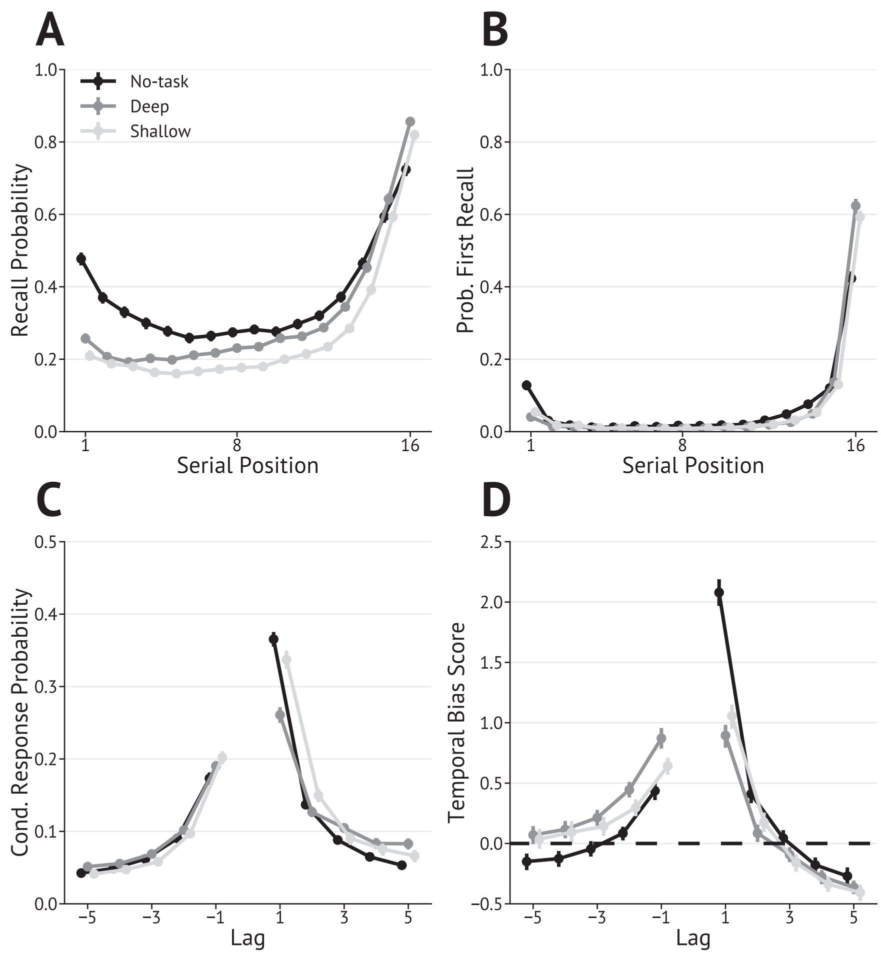

Figure 3.2 (A) Serial position curves, (B) probability of first recall curves, (C) lag-conditional response probabilities (lag-CRPs), and (D) temporal bias scores for no-task, deep processing, and shallow processing lists. Temporal bias scores for each lag were calculated by comparing the number of times a transition of that lag was actually made to the number of times it would be expected to occur by chance. Chance was calculated by permuting the order of recalls for each list 500 times and counting on average how many times each lag occurred for each permutation. The dotted line for the temporal bias scores indicates a score of zero (no bias). Error bars are bootstrapped 95% confidence intervals.

Lag-CRPs visualize the TCE as the conditional probability of making a transition of a given *lag*. Lag-CRPs displayed higher probabilities for near than far lags for all conditions (Figure 3.2C). The peak of the curve was largest for no-task and smallest for deep processing (cf. chance-adjusted TF scores in Figure 3.1B). While the no-task and shallow conditions exhibited the forward asymmetry typically associated with the TCE, this asymmetry was attenuated in the deep condition. However, caution is warranted in interpreting these results. Serial position effects can introduce a spurious TCE that disguises true differences between conditions, particularly when recall or primacy/recency differ substantially among conditions (Healey et al., 2019; Mundorf et al., 2021; Polyn et al., 2011; Uitvlugt & Healey, 2019), as they do here.

This spurious TCE can be illustrated by simulating data where items are recalled with no true TCE. To demonstrate, recalls were simulated for 100,000 participants for each condition (displayed in Figure 3.3). The probability of recalling each item was set to the recall probability of the corresponding position in that condition's serial position curve. This resulted in n items recalled for each simulated participant. To simulate data with *no* contiguity, the items' output order was randomly shuffled. Yet, the lag-CRPs (Figure 3.3B), still display a TCE with forward asymmetry. These lag-CRPs are heavily influenced by recency; $lag = +1$ is highest for shallow processing, the condition with the most recency. In contrast, temporal bias curves and chance-adjusted TF scores (Figure 3.3C and D) accurately display a null TCE for all conditions, making them better tools for comparing across conditions.

Returning to the data, temporal bias scores (Figure 3.2D) were highest for no-task, particularly at $lag = +1$. Forward asymmetry was reduced for shallow and completely eliminated for deep processing. Temporal bias scores reveal the higher TCE for deep processing (see Figure 3.1B) is due to the symmetrically high bias for near transitions, which results in overall greater temporal contiguity than the asymmetrical shallow condition.

Exploratory Followup Analyses

While there was a significant LOP effect on temporal contiguity, the effect was small. One possible explanation for the small effect size is that deep processing may also enhance semantic contiguity. Deep processing is inherently semantic and increases semantic organization, at least in lists with a category structure (e.g., Einstein & Hunt, 1980; Koriat & Melkman, 1987). However, items can only be recalled in one order. When items are presented in random order, organizing recalls by semantic similarity inherently reduces temporal contiguity. Thus, the LOP effect on the

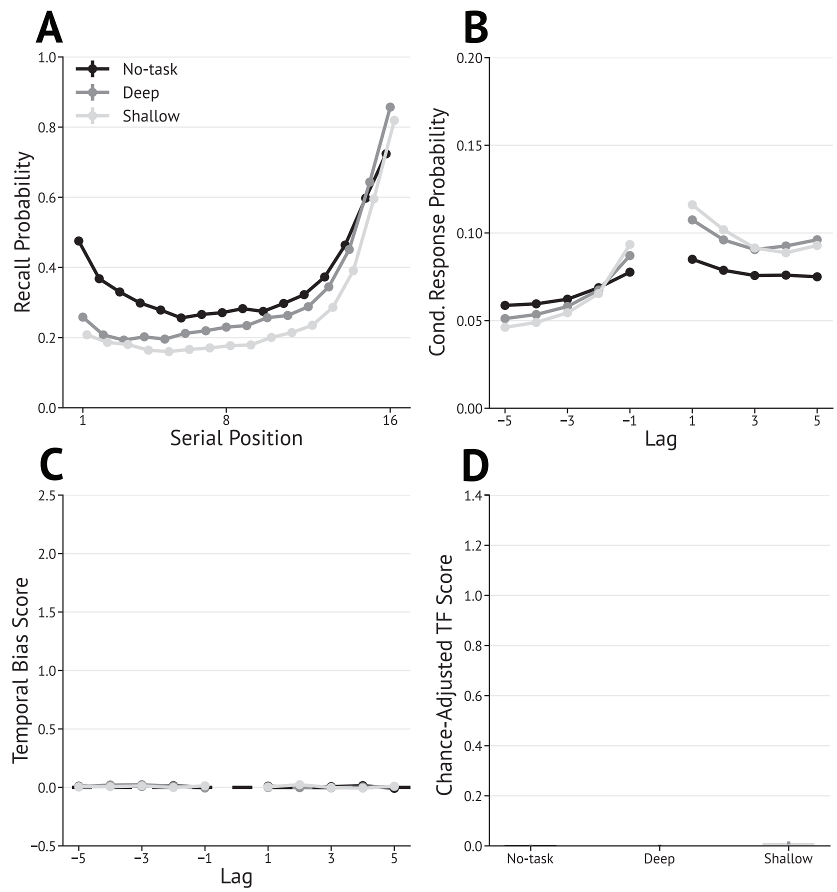

Figure 3.3 Simulated (A) serial position curves, (B) lag-conditional response probabilities (lag-CRPs), (C) temporal bias curves, and (D) chance-adjusted temporal factor (TF) scores from a model where recall order was randomly selected with regard to lag to produce simulated recalls with no temporal contiguity. Recalls were generated for 100,000 simulated participants, each recalling from 1 list of 16 items. For each participant, which items would be recalled was determined using binomial distribution where the probability of the participant recalling an item from a given serial position was set to the recall probability of the corresponding serial position in that condition's serial position curve. This resulted in *n* recalled items. Recall order was determined by randomly shuffling the *n* recalled items. Despite the data being generated such that items were recalled in random order (with zero temporal contiguity), the lag-CRPs display a contiguity effect as an artifact of the recency in the simulated serial position curves. Simulated lag-CRPs are presented on a smaller scale here in order to better display differences between conditions in this simulated data. Temporal bias curves display a null TCE, consistent with the method of data simulation. Chance-adjusted TF scores are also at or near zero for all conditions (making them barely visible in this figure).

TCE may have been attenuated by greater semantic organization in the deep condition.

Split-half Reliability for Individual Difference Variables

Table 3.1 Split-half reliability for recall probability, chance-adjusted temporal factor (TF) scores, and chance-adjusted semantic factor (SF) scores are presented here. For each condition, split-half reliability was calculated following the methodology of Sederberg et al. (2010). For each participant, I stratified their valid lists (where at least 2 list items were recalled) by condition and then randomly divided the participant's lists into two sets. In cases where the participant had an uneven number of valid lists in a given condition due to exclusions, I randomly selected which set would contain an additional list for that participant. I calculated probability of recall and chance-adjusted factor scores for each set and correlated the scores for set 1 with scores for set 2, correcting with the Spearman-Brown prediction formula ($2\rho/[1+\rho]$). This procedure was repeated 2,000 times, where the lists assigned to each set were randomly chosen for each participant in each iteration.

Condition	Recall prob.	Chance-adjusted TF scores	Chance-adjusted SF scores
No-task	0.923	0.759	0.072
Deep	0.892	0.628	-0.013
Shallow	0.897	0.671	0.058

Semantic Contiguity

In all conditions, chance-adjusted SF scores were small but above chance (Figure 3.1C). A repeated measures ANOVA revealed a significant effect of condition on semantic contiguity, $F(2, 1358) = 5.21$, $p = .006$, $\eta^2_G = .005$. Post-hoc tests with a Bonferroni adjusted[1] $\alpha = .006$ revealed greater semantic contiguity in the no-task ($M = 0.10$, $SE = 0.01$) compared to the shallow condition ($M = 0.05$, $SE = 0.01$), $t(679) = 3.34$, $p = .001$, $d = 0.117$. There were no differences between no-task and deep ($M = 0.06$, $SE = 0.01$), $t(679) = 2.55$, $p = .011$, or deep and shallow, $t(679) = 0.61$, $p = 0.540$.

Individual Differences

Individual differences in recall, temporal contiguity, and semantic contiguity were also considered. Reliabilities for recall and the chance adjusted factor scores are reported in Table 3.1. While recall and TF scores were fairly reliable, SF scores were quite unreliable in all conditions. Thus, correlations involving semantic contiguity are not reported.

The TCE was positively correlated with recall in no-task ($r(678) = .76$, $p < .001$), deep

[1] Adjusted α is .05/9. I conducted nine post-hoc analyses: three t-tests for semantic contiguity, three correlations between TCE and recall, and three correlations between semantic contiguity and recall.

($r(678) = .65$, $p < .001$), and shallow ($r(678) = .72$, $p < .001$) lists with a Bonferroni adjusted $\alpha = .006$, consistent with previous research using unrelated items (Mundorf et al., 2021; Sederberg et al., 2010; Uitvlugt & Healey, 2019).

Interim Discussion

I tested three hypotheses for how levels of processing (LOP; deep, shallow, no-task control) should influence the temporal contiguity effect (TCE). The first hypothesis was if deeper processing causes context to drift farther, the TCE should be greater for deep than shallow processing. The second hypothesis was if deeper processing instead increases processing of item information at the expense of order information, it should reduce the TCE. The final hypothesis was if the TCE arises from strategic control processes, *any* encoding task should disrupt it, regardless of depth. Both recall and the TCE were highest with no imposed processing task, were reduced under deep processing, and were further reduced under shallow processing. These results are inconsistent with the hypothesis that deep processing improves memory for items at the expense of memory for order. Instead, they support the hypothesis that deeper processing induces more context drift and the hypothesis that any imposed encoding task disrupts strategic processing.

Item-Order Account

These results are incompatible with the item-order account, which assumes any manipulation that draws attention to item-specific processing will reduce relational processing. If deeper processing enhances item-specific processing (Eysenck, 1979; Healey & Kahana, 2016), the TCE should be reduced. Yet, deeper processing *increased* the TCE. For the item-order account to be consistent with these results, major assumptions regarding the relationship between item and order information would have to change.

Retrieved Context Models

Under retrieved context models, the size of the TCE should depend on the degree to which context changes with each item studied. During encoding, context drift is driven by the items themselves activating their existing associations, and context travels farther during encoding when items strongly activate their pre-existing context. If context drifts farther with each item, only items studied nearby in time become associated with similar contexts. This should result in a greater preference for recalling items originally experienced nearby in time together, and the TCE will be large. Context changes very little, however, if items weakly activate their associated context. If context drifts only a short distance, all items form associations with similar contexts, reducing the preference for recalling nearby items together—the TCE will be small. The finding of a greater TCE for deep than shallow processing is consistent with the hypothesis that deep processing should activate more of items' associated contexts (Healey & Kahana, 2016), causing context to drift farther. Notably, these results are inconsistent with an alternate version of these models where deep processing acts only on the feature layer.

Although this hypothesis was framed purely in terms of whether the TCE was larger or smaller in deep processing and not the size of that difference, it is worth noting the observed effect size for the difference in temporal contiguity between deep and shallow processing was much smaller (d = .110) than might be expected. A small effect is still compatible with retrieved context models, where the change in context drift can be large or small, creating a larger or smaller TCE. The small effect does, however, suggest that large increases in temporal contiguity are not necessary for the beneficial effects of deep processing on memory, and temporal contiguity is only a part of the LOP puzzle.

Influence of Control Processes

These results are also consistent with accounts that assume the TCE arises from order-based encoding strategies. A strategic control processes account predicts that assigning any task during encoding will interfere with order-based strategies and thus reduce the TCE. Consistent with this prediction, recall and the TCE were greatest with no task. A strategic control processes account also predicts that, since order-based strategies encourage forward transitions, forward asymmetry should be greatest when no task interferes with strategy use. Indeed, asymmetry was greatest for the no-task condition.

Differences in strategy use may also provide an explanation for differences in asymmetry among the processing conditions. If participants have limited time to study, a more time-consuming task will leave less time for order-based strategies and result in less forward asymmetry compared to a shorter task. Thus, the reduced forward asymmetry for deep compared to shallow processing could be a result of the deep task taking longer to complete.

To test the plausibility of this explanation, I examined response times for participants' keypresses on the encoding task (how quickly they pressed "Y" or "N" in response to the shallow or deep processing question). Fifteen participants who made zero total responses were excluded. For each included participant, I calculated the mean response time for each processing task condition (deep, shallow). Supporting the explanation that participants had less time to engage in strategic control processes in deep versus shallow lists, participants responded more slowly to the deep ($M = 0.73$ s) than the shallow ($M = 0.68$ s) processing task, $t(644) = 31.44$, $p < .001$. While retrieved context models may be able to explain differences in asymmetry with existing mechanisms, the strategic control processes account offers a clear explanation for differences in asymmetry.

Conclusions

Recall and the TCE were higher under deep than shallow processing and highest with no encoding task. Retrieved context models and a strategic control processes account are each consistent with a subset of these results. Although theories based on context drift and those which emphasize strategy have been presented as conflicting explanations (Healey et al., 2019; Hintzman, 2016), integrating these two accounts provides the most comprehensive explanation of these results. Integrating a strategy account with retrieved context models will support development of a theory with well-defined mechanisms (even for the difficult-to-define strategic control processes) that accounts for both automatic and intentional processes of memory.

CHAPTER 4

DOES ORGANIZATIONAL STRATEGY DURING ENCODING DETERMINE LATER RECALL ORGANIZATION?

Strategic Control Processes at Encoding and Retrieval

The results of the previous experiment suggest that both the automatic TCE-generating mechanisms proposed by retrieved context models and strategic control processes are important to explain patterns of recall organization, and both perspectives should be considered in furthering theory development. One open question, however, is to what degree control processes during encoding determine the kinds of associations that are later available to guide recall. The answer to this question has theoretical importance; if temporal information is always automatically encoded, as predicted by retrieved context models, differences in recall organization would depend primarily on differences in how associative information is utilized at *retrieval* rather than the degree to which it is initially encoded. In contrast, if strategies during study control the encoding of associative information, then the TCE should be greatly reduced when participants focus on other kinds of associations during encoding. Because we cannot directly observe strategy use in most situations, it is unclear if variations in the TCE due to strategic control processes are a result of control processes implemented at encoding, retrieval, or both. To address this question, the effects of strategic control processes at encoding must be dissociated from the effects of retrieval strategies.

In free recall, patterns of recall organization can provide some insight into the kinds of strategic control processes participants adopt. For example, in lists of unrelated words participants often report using order-based strategies such as linking items together to form a story (Bouffard et al., 2018; Delaney & Knowles, 2005; Hintzman, 2016; Unsworth, 2016), and a robust TCE is observed. In lists with semantic structure, participants often take advantage of the available semantic associations, recalling semantically related words together, and the TCE is reduced (Healey & Uitvlugt, 2019; Hong et al., 2022; Polyn et al., 2011). The reliance on semantic rather than temporal organization could be a result of processes operating at either encoding or retrieval. The TCE could be reduced in lists with semantic structure because participants focus on semantic associations during encoding, leading to less encoding of temporal information; alternatively, a semantic strategy could come into play only at retrieval, resulting in memory search being guided primarily by semantic associations even though temporal information is available.

Directly manipulating participants' recall organization by assigning recall strategies can make the effects of control processes easier to distinguish. Healey and Uitvlugt (2019) found that instructions to focus on either semantic or temporal associations had a significant impact on later recall organization. In this experiment, participants studied either related lists (lists composed of several clusters of semantically related words randomly shuffled throughout the list) or unrelated lists. Prior to encoding, participants were instructed to either try to recall items in temporal order (temporal-focus), try to recall semantically related items together (meaning-focus), or use whatever strategy they preferred (free recall). The TCE was reduced by meaning-focus instructions and was completely eliminated in the meaning-focus condition for related lists. Additional analyses revealed a small TCE for transitions within a semantic cluster even in the absence of an overall effect, suggesting that some temporal information is automatically encoded. These findings provide strong evidence that strategic control processes modulate the TCE.

Yet, even when organizational strategies are directly manipulated, the effects of strategic control processes at encoding and retrieval are still not easily distinguished. The goal of Healey and Uitvlugt's (2019) experiment was to examine the effects of different strategies overall, so their instructions to adopt a temporal or semantic focus were given to participants prior to encoding the first list. As a result, strategic control processes at either encoding or retrieval could be responsible for the observed differences in temporal and semantic organization.

During encoding, strategic control processes could be used to prioritize task-relevant associations. In this way, encoding strategies could determine what kinds of information are later available to guide retrieval. Since attentional resources during encoding are limited (Kahneman, 1973), focusing on only task-relevant information is generally adaptive and should improve memory performance (Benjamin, 2007). Strategic control processes could thus determine the *amount* of temporal information that is encoded (although the encoding of temporal information cannot be entirely attributed to encoding strategies, see Chapter 2; Mundorf et al., 2021). The finding that the TCE is greatly reduced when participants are instructed to adopt a semantic retrieval strategy and ignore temporal order (Healey & Uitvlugt, 2019) supports this possibility. Even a simple semantic processing task during encoding ("Is this item alive?") can reduce the TCE (Long & Kahana, 2017), perhaps by making semantic associations more task-relevant and leading to less encoding of temporal information. If the differences in organization observed by Healey and Uitvlugt (2019) are primarily due to control processes at encoding, then a semantic encoding strategy should interfere with participants' ability to organize items later based on a different kind of association, regardless of their retrieval strategy.

Another possibility is that temporal information is always automatically encoded to the same extent, but strategic control processes at retrieval determine how those associations are used to guide recall. The basic mechanisms which allow for memory formation and retrieval in retrieved context models naturally give rise to a TCE, so these models assume temporal information is automatically encoded whenever memories are formed. Therefore, retrieved context models predict that even though a semantic retrieval strategy may reduce the TCE, temporal information should still be *available* to guide recall. These models also predict the availability of semantic information during recall, albeit through different mechanisms than those responsible for temporal contiguity. In most implementations of retrieved context models, semantic information is represented in the form of pre-existing knowledge. Because these associations are formed prior to study, they should also be available regardless of participants' encoding strategies, and the amount of semantic contiguity should depend on control processes at retrieval.

The goal of the present study is to examine the effects of temporal and semantic organizational strategies when participants are either assigned the same strategy for both encoding and recall or when they are forced to switch strategies at recall. If there is a trade-off between the encoding of temporal and semantic information, then focusing on one dimension during encoding should impair participants' ability to organize along the other dimension at retrieval (Long & Kahana, 2017). In contrast, information that is automatically encoded should be available at retrieval regardless of encoding strategy. If temporal information is encoded automatically, participants should be able to use it when they are instructed to adopt a temporal retrieval strategy, even if during encoding they were instructed to focus on semantic associations and ignore temporal order. A similar pattern would be expected for semantic associations.

General Methods

General Materials

Two lists of 16 words each were created for each participant. These words were drawn from a pool of 1,638 nouns developed for PEERS. The items for each list were selected such that semantically related words were located at both near and far temporal lags so that some semantic relationships would be present in the list but temporal and semantic organization would not be in complete opposition or complete congruence with one another. Lists were created following a similar procedure to that used in the original PEERS study (for a description, see Healey et al., 2019). Semantic similarity was measured using the Word Association Space (WAS) model (Steyvers et al., 2004). Each list was composed of eight pairs of words, where two pairs were drawn from the following four similarity bins: high similarity ($\cos\theta$ between words > 0.7), medium-high similarity ($0.4 < \cos\theta < 0.7$), medium-low similarity ($0.14 < \cos\theta < 0.4$), or low similarity ($\cos\theta < 0.14$). The lists were arranged such that one pair from each bin was presented in adjacent serial positions and the other pair was separated by at least two other items.

General Procedure

Because participants' understanding of the instructions at encoding and retrieval was critical for this study, two pilot studies were first conducted to test participants' ability to follow the instructions. Both pilot studies and Experiment 3 followed the same procedure except where otherwise noted.

Each participant studied 2 unique lists of 16 items each. Before studying the first list, participants were instructed about two different recall strategies (temporal and semantic) and provided examples of each. Participants' understanding of these two strategies was tested with two multiple-choice comprehension checks. For comprehension checks 1 and 2, an example list was presented

(TOWN, BUNNY, SHOP, CRICKET, WHISKERS, CAFE). To test if participants understood the temporal strategy, they were required to answer the following question: “Assume the first word that popped into your mind was BUNNY. Which word might you next recall if you were using the ORDER in which the words were originally presented to guide your memory search? Press the key corresponding to the correct option.” The answer options were 1) CAFE, 2) WHISKERS, 3) SHOP. The correct answer, SHOP, was always presented as option 3. To test if participants understood the semantic strategy, they were presented with the same list again (TOWN, BUNNY, SHOP, CRICKET, WHISKERS, CAFE) and asked the following question: “Assume the first word that popped into your mind was BUNNY. Which word might you next recall if you were using the SEMANTIC relationships between the words to guide your memory search? Press the key corresponding to the correct option.” The answer options were 1) CAFE, 2) SHOP, 3) WHISKERS. The correct answer, WHISKERS, was always presented as option 3. These are referred to as comprehension check 1 (temporal) and comprehension check 2 (semantic).

Then, depending on their condition, participants were instructed to either “try to use only the original order of the list to guide your memory search” (initial temporal focus) or “try to use only the semantic associations of items in the list to guide your memory search” (initial semantic focus) on the memory test that would follow each list. The first list served as practice. During study, each word was presented for 1.5 s with a 500 ms inter-stimulus interval. After studying the first list, participants completed 16 s of a math distractor task. They were then allowed 60 s to recall as many words as they could remember from the list. Participants recalled items one at a time by typing them into an empty text box, pressing ENTER after each word. A reminder of their assigned retrieval strategy appeared onscreen during the entire recall period.

Before studying the second list, participants were presented with a brief reminder of their assigned retrieval strategy. They then studied List 2 using the same encoding strategy as before. The study portion of List 2 used a new list of words but was otherwise identical to the study period for List 1. The critical manipulation took place before recall of List 2. Depending on their condition, participants were either instructed to use the *same* strategy as they were instructed to use during List 1 or a *different* strategy. In the congruent temporal/temporal and semantic/semantic conditions (and free/free in Pilot Study 1), participants were simply reminded of their assigned strategy. In the incongruent (semantic/temporal and temporal/semantic) conditions, participants were instructed to use a different strategy than they were previously assigned. Participants in the temporal/semantic condition received these instructions immediately before recalling List 2: "When you originally studied this list, you focused on the order of the words. But now, we want you to try to use ONLY the SEMANTIC (meaningful) associations between the words to guide your memory search. Completely ignore the order in which you saw the words." These instructions were followed by an example of a semantic recall strategy. Participants in the semantic/temporal condition received these instructions: "When you originally studied this list, you focused on the semantic (meaningful) relationships between the words. But now, we want you to remember the words based ONLY on the ORDER that you originally studied them. Completely ignore any semantic relationships." Although the instructions were self-paced, the number of words in the instructions for each condition was approximately the same, resulting in a similar delay between study and recall for all conditions.

To ensure participants understood which strategy they were being asked to use for recalling List 2, they were required to complete one final multiple-choice question (comprehension check 3), where they selected which recall strategy they were supposed to use on the upcoming memory test. The four possible response options were: 1) "Recall the words based ONLY on the SEMANTIC relationships between them," 2) "Recall the words in whatever order they come to mind," 3) "Recall the words based ONLY on how similar they SOUND when said aloud," and 4) "Recall the words based ONLY on the ORDER that you originally studied them." The order of the answer choices varied, but the correct answer was always the last option presented. Participants indicated their answer by pressing the corresponding key. The List 2 recall period was identical to recall for List 1. After completing recall of List 2, participants completed a short post-survey questionnaire where they provided demographic information and reported encoding and retrieval strategies for each list. Participants were also asked to report if there was any reason to exclude their data (e.g., a large interruption, cheating by writing the words down as they appeared). On average, the entire task took less than 12 minutes.

The pilot studies and Experiment 3 differed in the exact conditions participants were assigned to (see details below).

Pilot Study 1: MSU Undergraduates Online

The goal of the first pilot study was to test participants' ability to understand and follow the instructions to focus on either temporal or semantic associations during encoding and understand when those instructions changed at retrieval. Data collection occurred online on participants' personal computers. A total of 338 Michigan State undergraduate students participated.

Participants were assigned to one of 6 conditions in a 3 (initial focus) x 3 (List 2 test focus) design. Prior to encoding List 1, participants were either instructed to use a temporal strategy (initial temporal focus), a semantic strategy (initial semantic focus), or to use whatever kind of strategy they preferred (free). The second manipulation took place before recall of List 2 where participants were instructed to either use a temporal strategy, semantic strategy, or to freely recall whatever words came to mind. Some participants received the same instructions for both encoding and retrieval (temporal/temporal, semantic/semantic, free/free conditions), while others received different instructions (temporal/semantic, temporal/free, semantic/temporal, semantic/free, free/temporal, free/semantic conditions).

Pilot Study 1 Results

The purpose of Pilot Study 1 was to determine if participants understood when they were instructed to switch strategies. This can be determined by considering performance on the comprehension checks. In particular, did participants pass the third comprehension check where they were required to indicate which retrieval strategy they were assigned to use for recall of List 2? Although 75% of participants correctly answered the critical comprehension check (check 3), only 42% of participants answered all three comprehension checks correctly. The low pass rate for the comprehension checks precluded any conclusions that could be drawn from the data.

Pilot Study 2: Paid Online Participants

The purpose of Pilot Study 2 was to see if a more highly motivated set of participants would be able to pass the comprehension checks at a higher rate. This pilot study was conducted on Prolific, an online data collection platform where participants were paid approximately $15.77 per hour. Because the sample size was constrained for this online paid study, there were only 2 conditions: temporal/temporal and semantic/temporal. Initially, participants were instructed to use either a temporal or semantic strategy (initial focus), but immediately before recall of List 2, all participants were instructed to use a temporal recall strategy (List 2 test focus). Otherwise, the procedure was identical to Pilot Study 1.

Performance in Pilot Study 2
Table 4.1 Average and standard deviation (SD) for recall probabilities for List 1 and List 2 of each condition in Pilot Study 2.

Condition	Avg. List 1 Recall Prob. (SD)	Avg. List 2 Recall Prob. (SD)
Temporal/Temporal	0.42(0.20)	0.47(0.22)
Semantic/Temporal	0.40(0.16)	0.36(0.19)

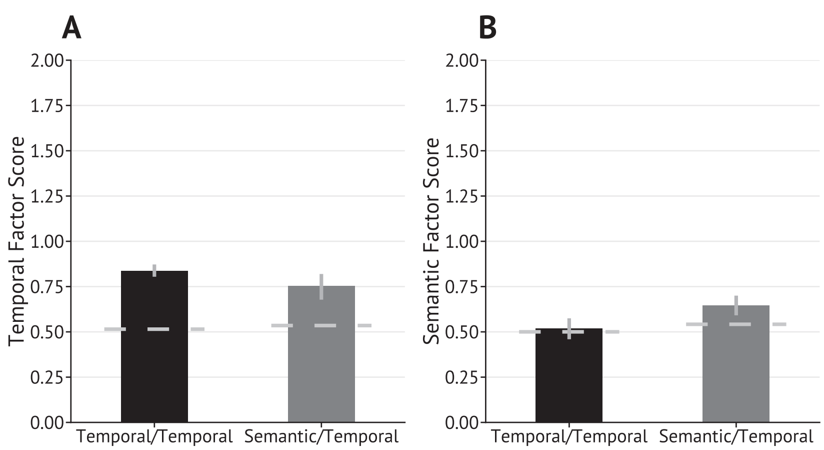

Figure 4.1 Measures of recall organization for Pilot Study 2 (List 2 recalls) are represented with (A) temporal factor (TF) and (B) semantic factor (SF) scores. For both TF and SF scores, chance was determined by permuting the order of recalls 5,000 times and calculating the average factor score of the distribution for each participant. Chance is indicated with a dotted line (calculated separately for each condition). Error bars represent bootstrapped 95% confidence intervals.

Pilot Study 2 Results

Performance on the comprehension checks in Pilot Study 2 was substantially higher than in Pilot Study 1. Of the 71 participants who completed the task, 96% passed comprehension check 3, which was the critical check for ensuring participants took note of the retrieval instruction manipulation, and 82% passed all three attention checks. Performance on attention check 3 was high regardless of whether participants had to switch strategies: no switch (temporal/temporal) = 97%, switch (semantic/temporal) = 93%.

When considering only those participants who correctly answered comprehension check 3, both conditions had similar levels of recall success on List 1 and List 2 (see Table 4.1). Both conditions also displayed significant temporal contiguity, as temporal factor (TF) scores for both conditions were significantly above chance (see Figure 4.1A; for a detailed description of TF scores see Chapter 2). Interestingly, although all participants were instructed to ignore semantic relationships during recall, participants who had studied with a semantic recall strategy in mind (semantic/temporal) still displayed a significant semantic contiguity effect. Semantic factor (SF) scores were significantly above chance in the semantic/temporal condition, but not in the temporal/temporal condition (see Figure 4.1B; for a detailed description of SF scores, see Chapter 3). One potential explanation for this is that although temporal information is automatically encoded, encoding strategies still influence later recall organization.

The results of Pilot Study 2 indicate that when sufficiently motivated, participants can understand the task instructions. They also provide some initial evidence that encoding strategies may affect the kind of information that is later used to guide retrieval.

Experiment 3

After establishing that sufficiently motivated participants were able to understand the instruction to switch strategies, I conducted a full-scale experiment to compare the effects of different encoding and retrieval strategies on recall organization.

Percentage Passing Attention Check 3 by Condition
Table 4.2 T/T = temporal/temporal, S/T = semantic/temporal, T/S = temporal/semantic, S/S = semantic/semantic.

Condition	T/T	S/T	T/S	S/S
Percent included	88.8%	82.5%	89.2%	94.6%
N included	153	139	151	162

Methods

Participants

Participants were Michigan State University undergraduate students who completed the experiment for course credit. A sample size of 400 (100 per condition) would provide 95% power to detect a main effect of strategy instructions with a moderate effect size ($\eta_p^2 = 0.13$). Although the effect of strategy instructions on recall organization reported by Healey and Uitvlugt (2019) was large ($\eta_p^2 > 0.3$), because the proposed study uses a different set of materials that do not set semantic and temporal contiguity in opposition to each other, the effect of strategy on organization may be smaller. Therefore, a sample large enough to detect a smaller effect was collected.

Pilot Study 1 demonstrated that undergraduate participants completing the study online from their own personal computers were unable to pass the comprehension checks at an acceptable rate. In-person undergraduates were expected to be comparable to the more highly motivated Prolific participants in Pilot Study 2 and it was expected that at least 80% would pass the critical comprehension check. To allow for anticipated exclusions, I collected data from additional participants beyond the target sample size. In total, 682 participants completed the experiment. Of these, 605 (88.7%) passed attention check 3 and were included in the analysis—a much higher rate than the online undergraduates in Pilot Study 1. The rates of passing comprehension check 3 were similar for all conditions, as displayed in Table 4.2. Of these included participants, 514 identified as female, and the mean age was 19.5 years ($SD = 3.5$).

Design

Participants were assigned to one of four experimental conditions in a 2 (initial focus: temporal, semantic) x 2 (List 2 test focus: temporal, semantic) between-subjects design. Some participants were assigned the same recall strategy throughout the experiment (temporal/temporal and semantic/semantic), while other participants' assigned strategy changed immediately before recall of List 2 (temporal/semantic and semantic/temporal).

Procedure

The procedure, task, and materials were identical to Pilot Studies 1 and 2 with two exceptions: Experiment 3 included the four experimental groups described above, and the experiment was conducted in person.

Results

List 1 Analyses

Although the primary focus of this experiment is how recall organization is affected by the change in strategy instructions for List 2, List 1 recalls are also informative insofar as they provide a baseline for the effect of the initial strategy instructions. For List 1, participants had only received their initial instructions, so the conditions with the same initial instructions (e.g., both the temporal/temporal and temporal/semantic groups) should display similar levels of recall.

To measure recall, I considered recall probabilities, serial position curves, and probability of first recall curves. Recall probabilities, which measure overall recall, provide a straightforward test of how initial instructions affected recall performance (Figure 4.2A). Initial focus did affect recall performance, $F(1, 601) = 48.74$, $p < .001$, $\eta^2 = .075$, and a semantic strategy produced overall higher recall, $t(603) = 6.92$, $p < .001$, $d = 0.563$. There was no effect of List 2 test focus, $F(1, 601) = 0.01$, $p = .936$, as would be expected since the testing instructions had not yet been provided, and no interaction, $F(1, 601) = 2.74$, $p = .098$. It is unsurprising that recall was better for participants who received the semantic focus instructions; when semantic similarities are present in a list, a semantic strategy tends to improve recall performance (Healey & Uitvlugt, 2019;

Mandler, 1967; Tulving & Pearlstone, 1966).

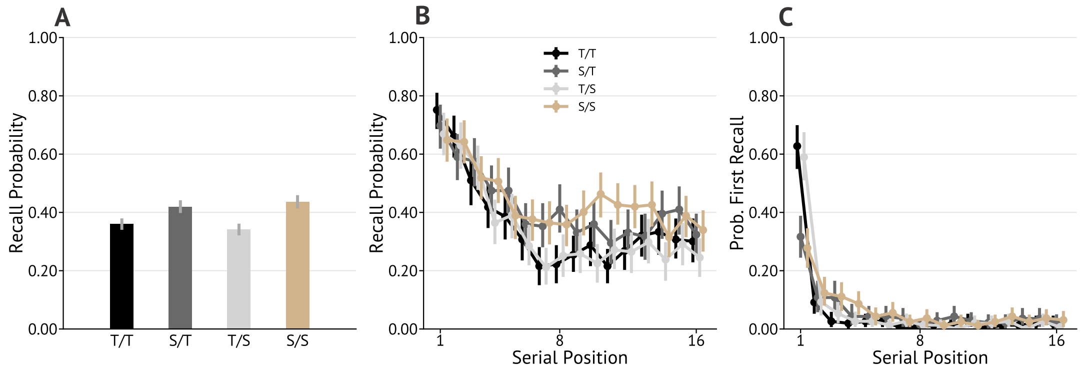

Figure 4.2 List 1 recall performance by condition, measured with (A) recall probabilities, (B) serial position curves, and (C) probability of first recall curves. Error bars represent bootstrapped 95% confidence intervals. T/T = temporal/temporal, S/T = semantic/temporal, T/S = temporal/semantic, S/S = semantic/semantic.

More detailed measures of recall, such as serial position curves and probability of first recall curves, reveal additional differences between conditions. There was a clear primacy effect in both overall recalls (serial position curves in Figure 4.2B) and in recall initiation (probability of first recall curves in Figure 4.2C). The tendency to initiate recall from the beginning of the list was much stronger for conditions initially instructed to use a temporal strategy. This is consistent with previous work where participants encouraged to recall items in temporal order tended to initiate recall from the beginning of the list (Healey & Uitvlugt, 2019).

The critical question of this study concerned how recall strategies affected recall organization. Temporal organization was examined using chance-adjusted TF scores and temporal bias scores (see Chapters 3 and 2 respectively for descriptions of these analyses). Chance-adjusted TF scores, displayed in Figure 4.3A, exceeded zero for all conditions. The temporal bias curves (Figure 4.3B) further illustrate a notable preference for near transitions, particularly transitions of *lag* = +1. Even when participants were instructed to ignore temporal order and focus on semantic similarities, temporal information still influenced recall order.

Semantic contiguity, measured using chance-adjusted SF scores, was also above chance in three of the four conditions: semantic/temporal, temporal/semantic, and semantic/semantic. Semantic organization above chance is indicated by chance-adjusted SF scores greater than zero in Figure 4.4B (see Chapter 3 for a description of how chance-adjusted SF scores are calculated). It is somewhat surprising that the temporal/semantic condition displayed semantic organization above chance while the temporal/temporal condition did not, given that both conditions received identical instructions for List 1. However, the difference between chance-adjusted scores for these two temporal conditions was not significant, $t(292) = -1.30$, $p = .196$, suggesting that this difference is fairly small and should not impact the interpretation of List 2 results.

More detailed patterns of semantic organization can also be considered by examining semantic lag-conditional response probabilities (semantic-CRPs). Semantic-CRPs calculate the probability of making a transition to another item of a given semantic lag, where a semantic $lag = 1$ represents making a transition to the item most semantically similar to the just-recalled item, and a $lag = 2$ represents making a transition to the second most semantically similar item, etc. This probability of making a transition of a certain lag is calculated conditional on an item at that semantic lag being available. The semantic-CRPs presented in Figure 4.4B provide further evidence that participants were more likely to make transitions among items that were highly semantically related. Across conditions, there was significant temporal contiguity in List 1 recalls, while semantic contiguity occurred primarily for participants who were instructed to use a semantic recall strategy.

List 2 Analyses

Recall Performance. Examining List 2 recalls allows for a test of how both initial focus (assigned before encoding) and test focus (assigned immediately before retrieval of List 2) influenced recall performance. Recall probabilities for each condition are presented in Figure 4.5A. There was a significant effect of initial focus, $F(1, 601) = 18.45$, $p < .001$, $\eta^2 = .029$, such that initially studying with a semantic focus resulted in higher recall than studying with a temporal focus, $t(680) = 4.17$, $p < .001$, $d = 0.336$. There was no effect of test focus on recall success, $F(1, 601) = 3.12$, $p = .078$. However, the effect of initial focus was qualified by an interaction, $F(1, 601) = 1.19$, $p < .001$, $\eta^2 = .019$.

As for List 1, a semantic strategy during encoding improved recall; however, this benefit was

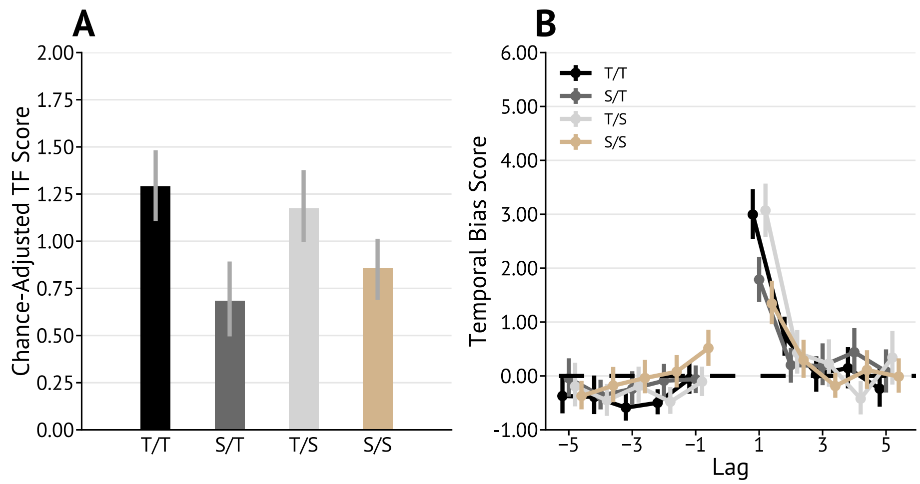

Figure 4.3 List 1 temporal contiguity by condition, measured with (A) chance-adjusted temporal factor (TF) scores, and (B) temporal bias scores. For TF scores, chance was determined by permuting the order of recalls 1,000 times. Chance-adjusted TF scores were calculated for each participant by subtracting the average of the chance distribution from the actual TF score for that participant and dividing by the standard deviation of the chance distribution. Temporal bias scores for each lag were calculated by comparing the number of times a transition of that lag was actually made to the number of times it would be expected to occur by chance. Chance was calculated by permuting the order of recalls for each list 1,000 times and counting on average how many times each lag occurred for each permutation. The dotted line for the temporal bias scores indicates a score of zero (no bias). Error bars represent bootstrapped 95% confidence intervals. T/T = temporal/temporal, S/T = semantic/temporal, T/S = temporal/semantic, S/S = semantic/semantic.

present in List 2 only for participants who also used a semantic strategy during recall (semantic/semantic condition), $F(1, 290) = 30.13$, $p < .001$. When participants switched to a temporal recall strategy (semantic/temporal condition), the benefits of using a semantic encoding strategy disappeared $F(1, 311) = 0.25$, $p = .617$. The higher recall for both the semantic/temporal and semantic/semantic conditions in List 1 may therefore have been due primarily to strategies at recall rather than encoding.

The serial position curves (Figure 4.5B) show recall was slightly higher for the semantic/semantic condition for mid-list items, but all conditions displayed a strong bias for recalling items from the beginning of the list as is typical for delayed free recall (Bjork & Whitten, 1974; Neath, 1993; Unsworth, 2008). The probability of first recall curves, presented in Figure 4.5C, also show a strong primacy effect. However, there was a clear effect of condition. The probability of

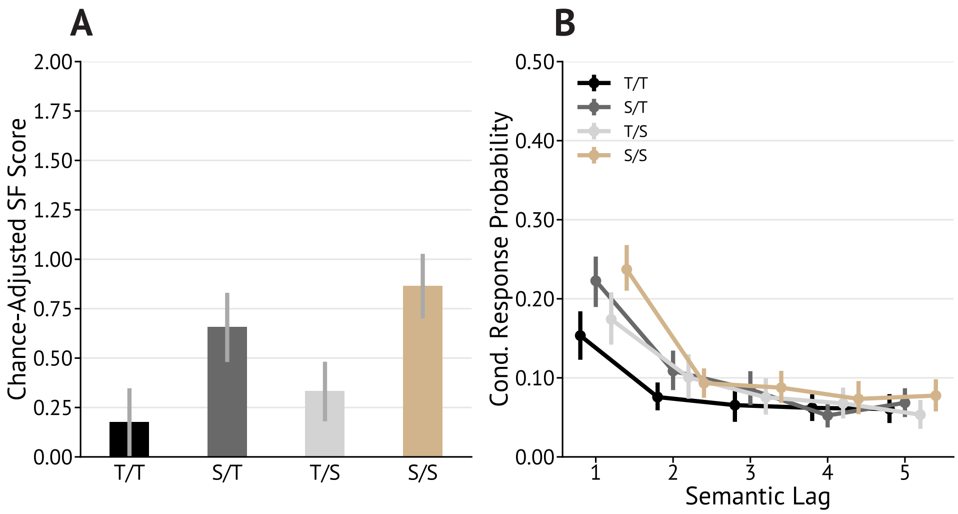

Figure 4.4 List 1 semantic contiguity by condition, measured with (A) chance-adjusted semantic factor (SF) scores, and (B) semantic lag-conditional probabilities (semantic-CRPs). For SF scores, chance was determined by permuting the order of recalls 1,000 times. Chance-adjusted SF scores were calculated for each participant by subtracting the average of the chance distribution from the actual SF score for that participant and then dividing by the standard deviation of the chance distribution. Error bars represent bootstrapped 95% confidence intervals. T/T = temporal/temporal, S/T = semantic/temporal, T/S = temporal/semantic, S/S = semantic/semantic.

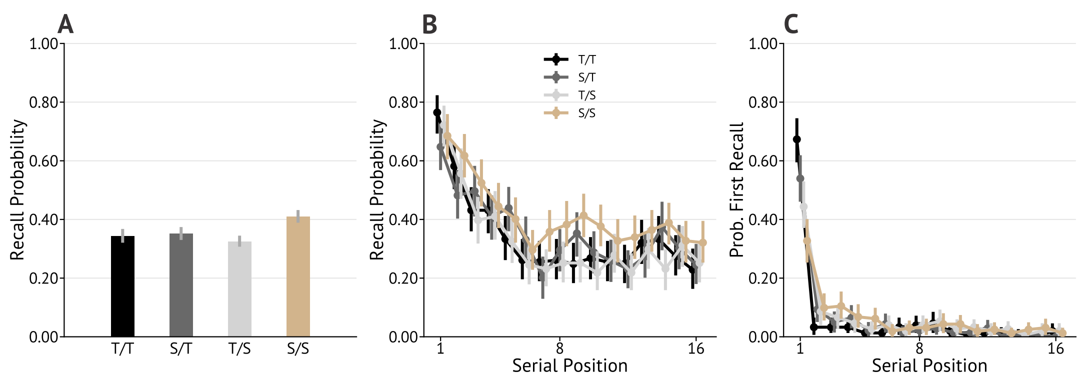

Figure 4.5 List 2 recall performance by condition, measured with (A) recall probabilities, (B) serial position curves, and (C) probability of first recall curves. Error bars represent bootstrapped 95% confidence intervals. T/T = temporal/temporal, S/T = semantic/temporal, T/S = temporal/semantic, S/S = semantic/semantic.

initiating recall from the beginning of the list was influenced by test focus (i.e., temporal/temporal and semantic/temporal have the two highest probabilities for serial position 1) with a smaller influence of initial strategy (i.e., probabilities of first recall for serial position 1 are higher for temporal/temporal versus semantic/temporal and for temporal/semantic versus semantic/semantic).

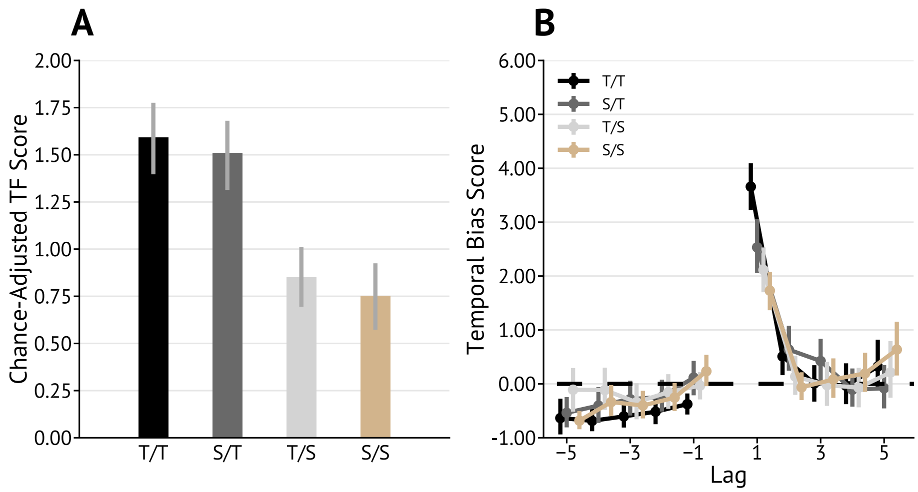

Figure 4.6 List 2 temporal contiguity by condition, measured with (A) chance-adjusted temporal factor (TF) scores and (B) temporal bias scores. For TF scores, chance was determined by permuting the order of recalls 1,000 times. Chance-adjusted TF scores were calculated for each participant by subtracting the average of the chance distribution from the actual TF score and dividing by the standard deviation of the chance distribution for that participant. Temporal bias scores for each lag were calculated by comparing the number of times a transition of that lag was actually made to the number of times it would be expected to occur by chance. Chance was calculated by permuting the order of recalls for each list 1,000 times and counting on average how many times each lag occurred for each permutation. The dotted line for temporal bias scores indicates a score of zero (no bias). Error bars represent bootstrapped 95% confidence intervals. T/T = temporal/temporal, S/T = semantic/temporal, T/S = temporal/semantic, S/S = semantic/semantic.

Temporal Contiguity. The main focus of this experiment was to investigate how task goals during encoding and retrieval influence recall organization, particularly temporal organization. Chance-adjusted TF scores exceeded zero in all conditions, supporting retrieved context models' prediction that temporal information is automatically encoded. Contiguity above chance is indicated by the 95% confidence intervals not crossing zero in Figure 4.6A.[1] Further, the TCE was not affected by initial focus, $F(1, 578) = 0.95$, $p = .329$. That is, participants' assigned strategy while they

[1]Calculating chance-adjusted TF scores requires participants to recall at least 2 list items in a given list so the chance distribution will have a standard deviation of more than zero. If at least 2 items are recalled, because positive and negative lags are treated as different values in this calculation, then at least 2 different TF scores are possible for each permutation. Chance-adjusted TF scores could not be calculated for the 21 participants who recalled fewer than 2 list items, so they were excluded from this analysis.

encoded List 2 did not determine the level of temporal contiguity in their recalls. However, there was a significant effect of *test focus*, $F(1, 578) = 65.15$, $p < .001$, $\eta^2 = .101$. Assigning a temporal retrieval strategy immediately before recalling List 2 resulted in a significantly greater TCE than a semantic retrieval strategy, $t(580) = 8.07$, $p < .001$, $d = 0.671$. There was no interaction between initial and test focus, $F(1, 578) = 0.006$, $p = .937$.

Temporal bias scores, presented in Figure 4.6B, also demonstrate an effect of test focus. However, these scores provide an additional insight: differences between conditions lie primarily in the bias for near forward lags (*lag* = +1). When examining this point in particular, there is some influence of initial focus during encoding. There is a much higher bias for *lag* = +1 transitions in the temporal/temporal condition, the condition with the highest levels of temporal contiguity. Bias for *lag* = +1 transitions is still high but lower for semantic/temporal, followed by temporal/temporal, and finally by semantic/semantic. Initial and test strategies had additive effects on this forward bias, with a greater influence of test focus on the final level of temporal contiguity.

Semantic Contiguity. Also of interest is how intentional strategies influenced semantic organization. Chance-adjusted SF scores,[2] displayed in Figure 4.7A, were influenced by test condition, $F(1, 575) = 16.91$, $p < .001$, $\eta^2 = .028$. As expected, participants with a semantic focus at test displayed greater semantic contiguity than those who did not, $t(577) = 4.11$, $p < .001$, $d = 0.343$. Semantic contiguity was not affected by study condition, $F(1, 575) = 2.87$, $p = .091$, and there was no interaction between study and test conditions, $F(1, 575) = 0.01$, $p = .926$.

The semantic-CRPs presented in Figure 4.7B provide a more detailed representation of how strongly participants were biased towards recalling highly semantically similar items together. The semantic/semantic condition was most likely to make transitions to the most semantically similar available item (semantic *lag* = 1). The temporal/temporal condition, where participants were instructed to ignore semantic associations at both encoding and at retrieval, was the least likely

[2]Calculating chance-adjusted SF scores requires participants to recall more than 2 list items for a given list so the chance distribution will have a standard deviation of more than zero. If 2 or fewer items are recalled, because semantic lags are only positive (in contrast to temporal lags, which can be both positive and negative), only one SF score is possible, no matter how the values are permuted. Chance-adjusted SF scores could not be calculated for the 23 participants who recalled 2 or fewer list items, so they were excluded from this analysis.

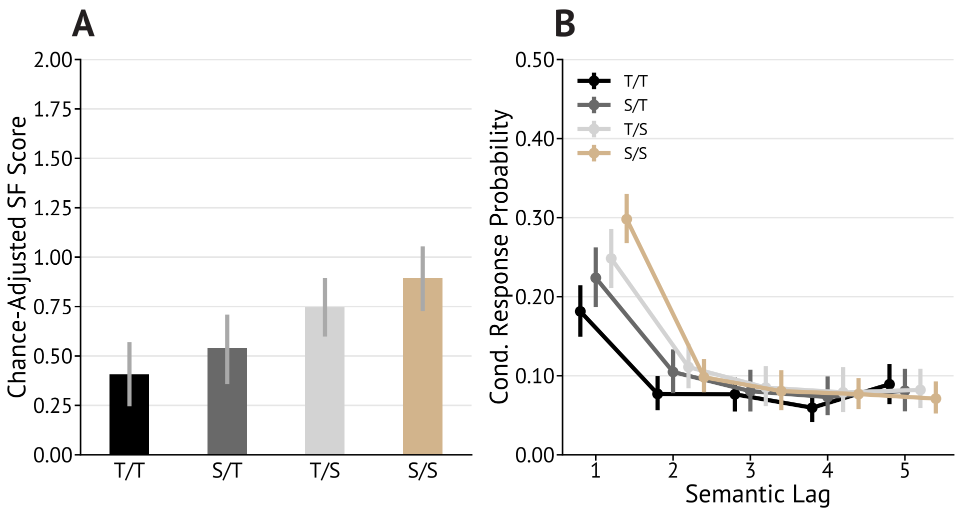

Figure 4.7 List 2 semantic contiguity by condition, measured with (A) chance-adjusted semantic factor (SF) scores, and (B) semantic lag-conditional probabilities (semantic-CRPs). For SF scores, chance was determined by permuting the order of recalls 1,000 times. Chance-adjusted SF scores were calculated for each participant by subtracting the average of the chance distribution from the actual SF score and then dividing by the standard deviation of the chance distribution for that participant. Error bars represent bootstrapped 95% confidence intervals. T/T = temporal/temporal, S/T = semantic/temporal, T/S = temporal/semantic, S/S = semantic/semantic.

to make transitions to items at a near semantic lag. Information about the semantic relationships between items may be readily available regardless of encoding strategy, since these relationships are based on prior knowledge. However, participants are able to selectively choose to utilize these associations to guide recall or not.

Individual Differences. Individual differences in recall, temporal contiguity, and semantic contiguity were also considered. Correlations between recall and both temporal and semantic contiguity are displayed in Figure 4.8. To account for multiple comparisons (8 correlations and 4 Fisher's z tests), a Bonferroni-adjusted $\alpha = .05/12 = 0.004$ was used for these comparisons.

Recall was positively correlated with temporal contiguity in all conditions. That is, temporal organization predicted recall success, regardless of participants' strategies during study or retrieval. The combination of initial and test strategies did, however, modulate the strength of this correlation. Correlations between recall and the TCE were significantly higher for temporal/temporal compared to both temporal/semantic, Fisher's $z = 4.19$, $p < .001$, and semantic/semantic, Fisher's $z = 4.81$,

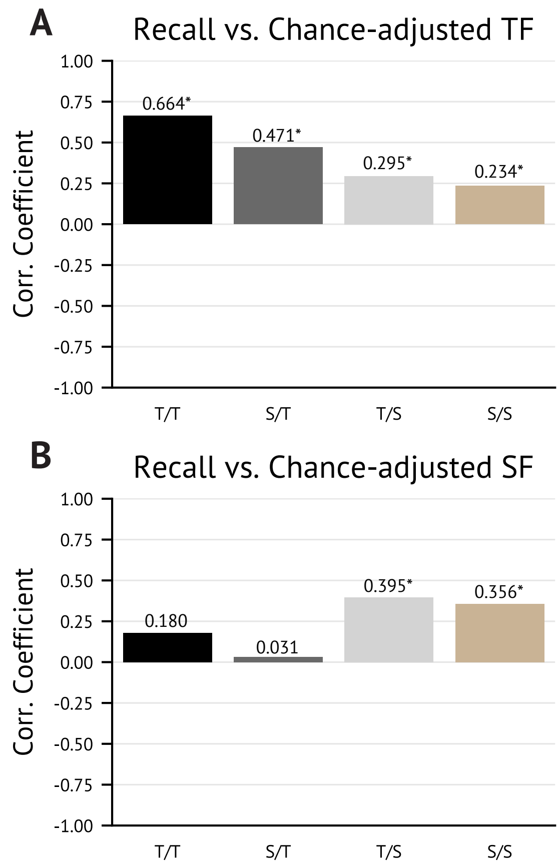

Figure 4.8 Correlations between recall and A) chance-adjusted temporal factor (TF) scores and B) chance-adjusted semantic factor (SF) scores for each condition. The displayed correlation coefficients represent Pearson's r, and correlation coefficients marked with * are significant at a Bonferroni-adjusted $\alpha = .004$. T/T = temporal/temporal, S/T = semantic/temporal, T/S = temporal/semantic, S/S = semantic/semantic.

$p < .001$. The TCE-recall correlation was also higher for temporal/temporal compared to the semantic/temporal condition, but this difference was not significant once the Bonferroni adjustment was applied, Fisher's $z = 2.38$, $p = .017$. Temporal contiguity predicted higher recall in all conditions, but this correlation was stronger when participants were assigned a temporal focus.

Semantic contiguity was also positively correlated with recall, but only in conditions assigned a semantic test strategy (the temporal/semantic and semantic/semantic conditions in Figure 4.8B), and there was no difference in the correlations between semantic contiguity and recall for the semantic retrieval conditions (Fisher's $z = 0.39$, $p = .693$). The finding that the task-relevant features are positively correlated with recall is consistent with Healey and Uitvlugt's (2019) findings for their temporal-focus (corresponding to the temporal/temporal condition here) and semantic-focus (corresponding to the semantic/semantic condition here) conditions, but they also suggest similar control processes could be responsible for participants' ability to maintain their assigned *recall* strategy and overall recall success.

Computational Modeling

Retrieved context models make the qualitative prediction that temporal information should be automatically encoded whenever memories are formed; therefore, a TCE should be observed in all conditions, even when participants were instructed to ignore temporal order during study. Fitting a computational model to the data, however, provides a quantitative test of whether retrieved context models are consistent with any trade-offs between temporal and semantic contiguity that might appear between conditions, particularly when participants are instructed to organize their recalls based on semantic similarity. To test this, a version of the Context Maintenance and Retrieval model (CMR) was fit to the data for each condition.

CMR is one of the family of retrieved context models that includes a representation not only of associations formed between items and their context during encoding but also pre-experimental semantic associations (meaningful associations between items formed prior to the experiment) and a representation of source context (used to represent task goals; Polyn et al., 2009a). In the past, this has allowed the model to explain patterns of both temporal and semantic organization, as well as clustering by encoding task. A detailed description of the model implementation used in this chapter, which largely follows Morton and Polyn's (2016) context-based semantic cuing model with a few exceptions, is provided in Appendix A. Table 4.3 provides a description of the parameters in this implementation of CMR.

In CMR, encoding proceeds as in the example presented in Figure 1.3. Episodic memories are formed when an association is established between a representation of the item being studied (represented on the feature layer in Figure 1.3) and the current state of temporal context (represented on the context layer in Figure 1.3). As each item is studied, its representation on the feature layer is activated, bringing to mind any pre-existing associations (pre-experimental context).

Following Morton and Polyn's (2016) context-based semantic cuing model, semantic associations were represented as part of the pre-experimental context associated with each item. Semantic associations were determined using Word Association Space (WAS; Steyvers et al., 2004), which represents each item as a vector in semantic space, and semantic similarity is measured as the cosine(θ) between them. The influence of these pre-existing semantic associations is controlled using the δ, α, and s_{cf} parameters listed in Table 4.3. These activated associations are then incorporated into the current state of context, causing context to change, or drift. When the next item is studied, it automatically forms an association with the current state of context, and it in turn activates its own pre-existing associations. Again, the newly activated pre-experimental context is incorporated into the current context, and context drifts. Importantly, when the next item is studied, the previous context representation is not completely erased; context is a blend of the previous context and the new context. As a result, items experienced relatively closer together in time are associated with more similar states of context. The context layer serves as a recency-weighted record of the past such that recent items are more strongly represented.

Immediately before the recall period, context drifts as two more events occur: the end-of-list distractor and the reinstatement of the beginning-of-list context. The rate at which context drifts for each event is controlled by the parameters $\beta_{distract}$ and β_{start} respectively (see Table 4.3). Then, during recall, context is used as a cue to retrieve items. When an item is recalled, both the pre-existing semantic associations of the just-recalled item and the context associated with the item during encoding are retrieved and incorporated into the current state of context. Semantically similar items provide good cues for one another in the same way as temporally adjacent items. They are associated with similar states of context, so when one item is retrieved, its associated context is a good cue for a similar item.

Methods

Although CMR has been shown to fit well to patterns of semantic and temporal contiguity in well-practiced participants studying lists of random words, it is an open question if the model will

Parameters for the Context Maintenance and Retrieval Model (CMR)

Table 4.3 Names and descriptions for each parameter used in the version of CMR implemented here.

Purpose	Parameter	Description
Encoding	ϕ_s	Scaling of primacy gradient in learning new context-feature associations
	ϕ_d	Rate of decay of primacy gradient
	γ_{fc}	Strength of new feature-context associations
	β_{enc}	Rate of context drift during encoding
	δ	Initial strength of pre-existing context-feature auto-associations
	α	Initial strength of other pre-existing context-feature associations
	s_{cf}	Semantic scaling parameter
Retrieval	β_{rec}	Rate of context drift during recall
	$\beta_{distract}$	Rate of context drift during the pre-recall distractor
	β_{start}	Amount of beginning-of-list context retrieved before recall begins
Decision Process	θ_s	Scaling of probability of recall failure
	θ_r	Rate at which the likelihood of recall success decreases with additional output position
	τ	Sensitivity to differences in activation at retrieval for luce choice rule

be able to explain the trade-offs in temporal and semantic contiguity observed here. The model has primarily been applied to free recall tasks where participants are able to adopt any strategy and temporal contiguity tends to be high. However, in Experiment 3, participants' goals had a strong influence on the degree of both temporal and semantic contiguity. To test the model's ability to explain levels of temporal and semantic contiguity across conditions, I attempted to fit the model to both the temporal bias scores and semantic-CRPs for each condition, as well as overall recall probability. I also fit the model to only the semantic-CRP, since CMR has not been directly fit to this measure of semantic contiguity before.

The model was fit to each condition 5 times for the fits to recall, temporal bias scores, and semantic-CRPs and 10 times for the fits to the semantic-CRPs alone using a genetic algorithm that ran for 5,000 generations with a total of $k = 13$ free parameters. Model fit was measured using root mean square deviation (RMSD). For details on the methods for model fitting and a table of best-fit parameter values, see Appendix A.

Results

Simulated data from the best-fitting parameter set for fits to the temporal bias scores, semantic-CRPs, and recall probabilities are presented alongside the behavioral data for each condition. Simulations for the temporal/temporal condition are presented in Figure 4.9. The best-fitting model was able to simultaneously fit the pattern of overall recall, temporal contiguity, and semantic contiguity in the temporal/temporal condition, particularly the strong tendency to make near forward transitions and the weak bias for making transitions of a semantic *lag* = 1. For the semantic/temporal (Figure 4.10), temporal/semantic (Figure 4.11), and semantic/semantic conditions (Figure 4.12), the model was also able to capture levels of temporal contiguity and recall but not the high levels of semantic contiguity.

The failure of the model to fit to the semantic-CRP for three out of the four conditions could be due to either a failure to capture the relationship between temporal and semantic contiguity *or* a failure to fit to high levels of semantic contiguity in general. To examine these possibilities further, I also fit the model to the semantic-CRP alone. Simulated data for fits to the semantic-CRP alone are presented in Figure 4.13. Although the best-fitting models were able to approximate the probability of making semantic *lag* = 1 transitions, none of the fits were exact. The best-fitting models for the semantic/temporal, temporal/semantic, and semantic/semantic conditions underestimated the probability of making semantic *lag* = 1. Not only that, but the 10 model fits for each condition were inconsistent. Fitting semantic contiguity was a challenge for this model, and perhaps there is only a small subset of the parameter space that can simulate semantic contiguity, particularly at high levels. Because it is not possible to cover the entire parameter space in 10 model fits, it is also possible that the model *could* fit to the high level of semantic contiguity in the semantic/semantic condition, given the right parameter set. However, the problem still stands: the model has substantial difficulty in fitting to high levels of semantic contiguity.

Interim Discussion

The aim of this experiment was to test if variation in the TCE due to strategy use is primarily

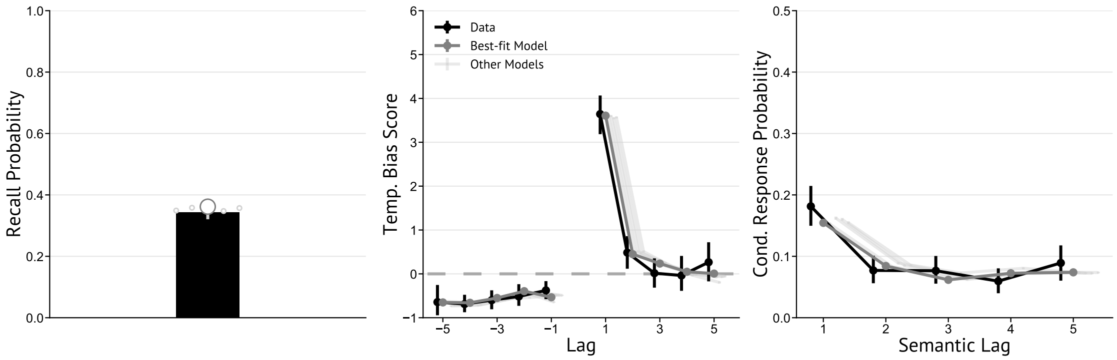

Figure 4.9 Behavioral data for the temporal/temporal condition alongside simulated recall probabilities, temporal bias scores, and semantic-conditional response probabilities (semantic-CRPs). Simulated data based on the best-fitting model of the Context Maintenance and Retrieval model are represented in dark gray. The best-fitting model was defined as the model with the lowest root mean squared deviation (RMSD) value. Additional model fits are plotted in light gray. Data from different model fits are offset slightly so similar fits can be distinguished. Error bars represent bootstrapped 95% confidence intervals.

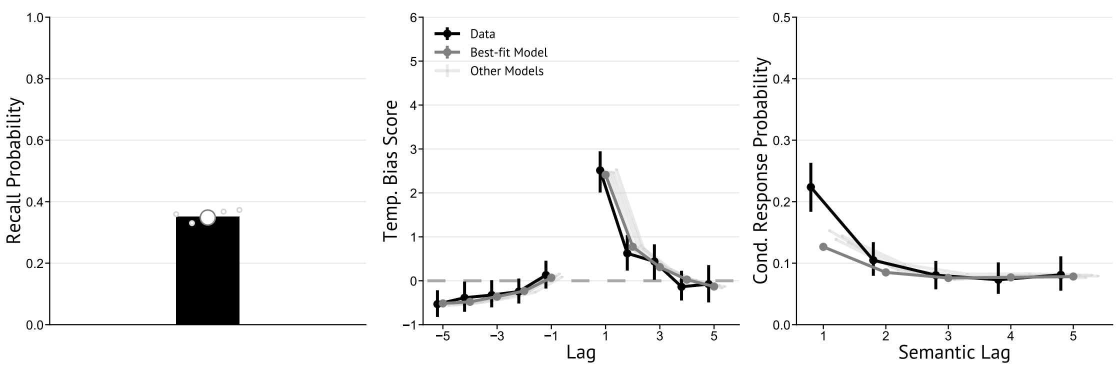

Figure 4.10 Behavioral data for the semantic/temporal condition alongside simulated recall probabilities, temporal bias scores, and semantic-conditional response probabilities (semantic-CRPs) from model fits to all three measures. Simulated data based on the best-fitting model of the Context Maintenance and Retrieval model are represented in dark gray. The best-fitting model was defined as the model with the lowest root mean squared deviation (RMSD) value. Additional model fits are plotted in light gray. Data from different model fits are offset slightly so similar fits can be distinguished. Error bars represent bootstrapped 95% confidence intervals.

a result of control processes at encoding or at retrieval. If control processes during encoding determine the kind of information that is later available, then focusing on semantic associations during encoding should reduce the TCE, even if participants are asked to focus on temporal order

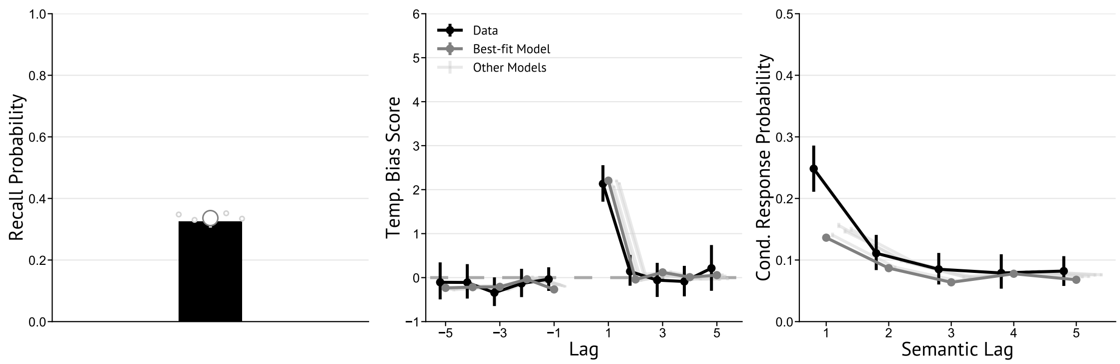

Figure 4.11 Behavioral data for the temporal/semantic condition alongside simulated recall probabilities, temporal bias scores, and semantic-conditional response probabilities (semantic-CRPs) from model fits to all three measures. Simulated data based on the best-fitting model of the Context Maintenance and Retrieval model are represented in dark gray. The best-fitting model was defined as the model with the lowest root mean squared deviation (RMSD) value. Additional model fits are plotted in light gray. Data from different model fits are offset slightly so similar fits can be distinguished. Error bars represent bootstrapped 95% confidence intervals.

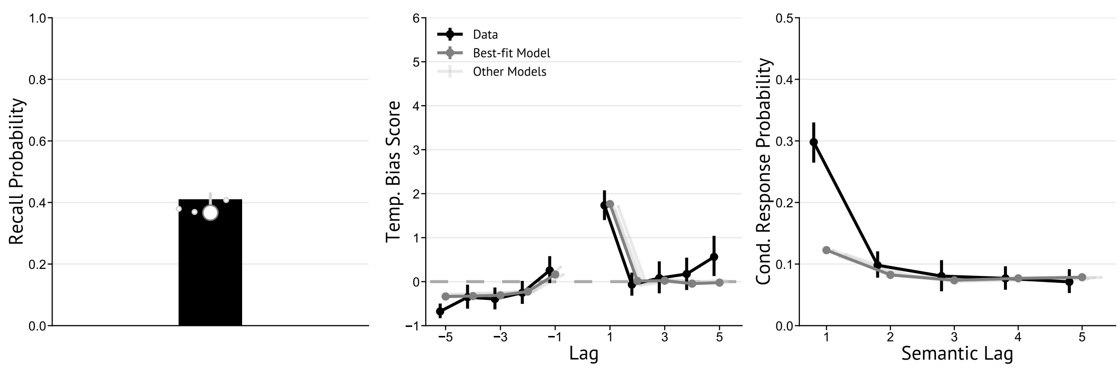

Figure 4.12 Behavioral data for the semantic/semantic condition alongside simulated recall probabilities, temporal bias scores, and semantic-conditional response probabilities (semantic-CRPs) from model fits to all three measures for the semantic/semantic condition. Simulated data based on the best-fitting model of the Context Maintenance and Retrieval model are represented in dark gray. The best-fitting model was defined as the model with the lowest root mean squared deviation (RMSD) value. Additional model fits are plotted in light gray. Data from different model fits are offset slightly so similar fits can be distinguished. Error bars represent bootstrapped 95% confidence intervals.

during retrieval. If strategic control processes at retrieval are primarily responsible for variations in the TCE and temporal information is always encoded automatically, then participants instructed to focus on temporal order at retrieval should display more temporal contiguity than those assigned

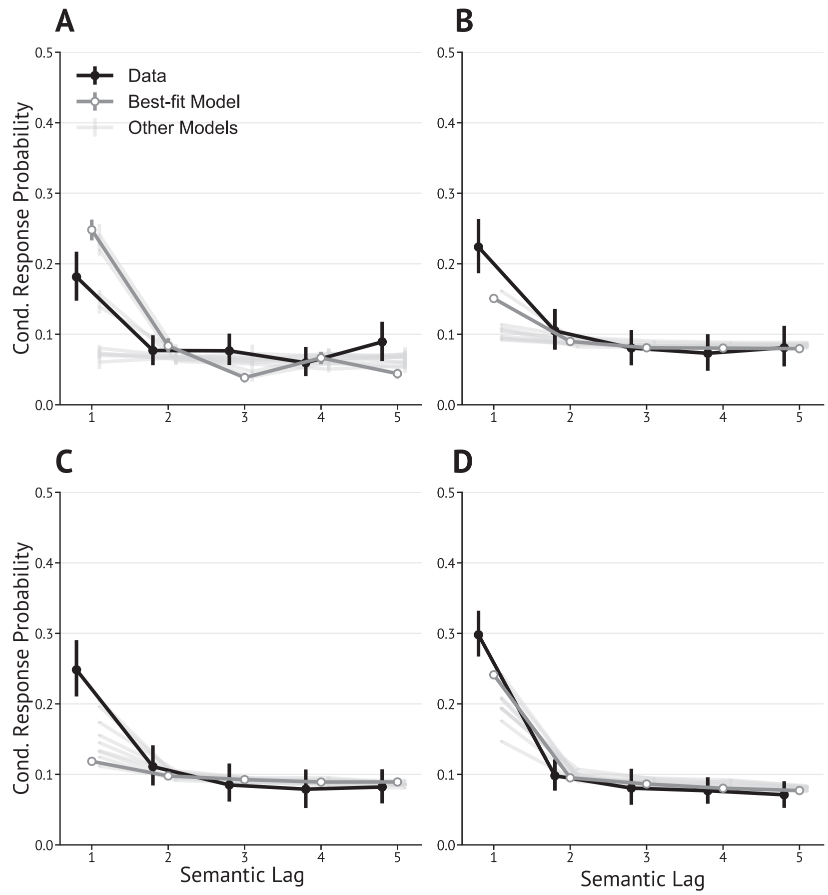

Figure 4.13 Behavioral data alongside simulated semantic-conditional response probabilities (semantic-CRPs) from model fits to the semantic-CRP only for A) the temporal/temporal condition, B) the semantic/temporal condition, C) the temporal/semantic condition, D) the semantic/semantic condition. Simulated data based on the best-fitting model of the Context Maintenance and Retrieval model are represented in dark gray with open points. The best-fitting model was defined as the model with the lowest root mean squared deviation (RMSD) value, even if it did did not provide the best fit to *lag*= 1. Additional model fits are plotted in light gray. Data from different model fits are offset slightly so similar fits can be distinguished. Error bars represent bootstrapped 95% confidence intervals.

a semantic focus at test, regardless of their initial strategy during encoding. Not only was temporal contiguity observed in all conditions, but there was also an effect of control processes at retrieval. A semantic focus at test greatly reduced temporal contiguity and increased semantic contiguity compared to a temporal focus at test. However, there was no effect of initial focus on either temporal or semantic organization. The context maintenance and retrieval model (CMR) was able to fit well to levels of overall recall and temporal contiguity although it was unable to fit to higher levels of semantic contiguity. These results are consistent with the account that attributes variations in temporal contiguity primarily to differences in intentional memory search at retrieval rather than encoding.

These results are not consistent with the hypothesis that strategic control processes at encoding determine the degree to which temporal information is learned. This encoding control processes account predicts that a semantic focus during encoding should interfere with the encoding of temporal information by directing participants' limited attentional resources to semantic associations instead. Therefore, a semantic focus during encoding should reduce, or even eliminate, the TCE. In this experiment, all participants displayed a TCE above chance even when they were instructed to ignore temporal order during encoding. This account could perhaps explain the significant TCE in all conditions by assuming that while *some* temporal order information is automatically encoded, control processes during encoding determine the *extent to which* temporal information is encoded. If so, the TCE should be larger for the initial temporal-focus conditions. Yet, further contradicting this account's predictions, participants' focus during encoding did not affect the TCE. A similar pattern was also present for semantic similarity. Participants in all conditions tended to group semantically related items together at recall, regardless of their initial focus.

The finding that task goals during encoding did not affect recall organization is somewhat surprising, as previous work has found that participants' ability to engage in control processes does affect temporal organization. For example, assigning a semantic processing task during encoding reduces the TCE (Chapter 3; Long & Kahana, 2017), and the TCE is greater under intentional versus incidental encoding (Healey, 2018; Mundorf et al., 2021). While it is true that the critical manipulation took place during encoding in these other experiments, participants' focus (or lack thereof) during encoding could have also influenced their approach to retrieval. This can also be seen in the present experiment by comparing recall dynamics for List 1 and List 2. In List 1, when all participants used the same strategies during both encoding and retrieval, the pattern was markedly similar to that observed in the parallel conditions in Healey and Uitvlugt (2019). The initial temporal focus conditions displayed high levels of temporal contiguity with low levels of semantic contiguity, and the initial semantic focus conditions displayed low levels of temporal contiguity with high levels of semantic contiguity. But in List 2, recall organization was affected only by participants' focus during retrieval. The effect of initial focus observed in List 1 could therefore also be attributed to participants' goals during retrieval.

These results are more consistent with an account that explains differences in recall organization as a result of automatic mechanisms operating at encoding and control processes operating at retrieval. Automatic encoding of temporal information is a core mechanism of retrieved context models, which posit that temporal information is recorded in the form of item-context associations whenever a new memory is formed. Therefore, these models predict a TCE even when participants do not strategically focus on temporal order during encoding. The results of Experiment 3 are consistent with this account. A TCE was observed even when participants were instructed to ignore temporal order information, and the TCE was unaffected by participants' focus during encoding (semantic or temporal). Temporal contiguity also predicted better recall in all conditions, even those instructed to ignore temporal order throughout the entire experiment, supporting the claim that memory for items and memory for their order are tightly linked together. These findings add to a growing body of work supporting the claim that temporal information is encoded automatically, even when participants are prevented from engaging in temporal encoding strategies (Chapter 2; Healey & Uitvlugt, 2019; Mundorf et al., 2021; Murphy & Castel, 2021).

Even if temporal information is automatically encoded, that only means that temporal information is available at recall, not necessarily that it *must* guide recall. In the search through memory space, many paths are possible. Strategic control processes during retrieval could be responsible for directing memory search to follow either temporal or semantic cues, or a combination of both, in response to task demands. Such control processes would be compatible with a retrieved context model explanation and are necessary to explain the present results. There was a clear effect of retrieval strategy on every measure of recall organization considered here, demonstrating participants were capable of exerting cognitive control over their memory search when instructed to do so.

Further, even where there was a small effect of participants' assigned focus during study, that effect was modulated by retrieval focus. Both initial and test focus influenced probability of recall initiation from the beginning of the list (probability of first recall in Figure 4.5C) and the bias for near forward lags (temporal bias scores in Figure 4.6B), two measures that are characteristic of a temporal strategy. Both of these measures were highest for the temporal/temporal condition, followed by semantic/temporal, temporal/semantic, and finally semantic/semantic. There was a small influence of initial focus, but the pattern was dominated by the effect of test focus; the two conditions with the highest scores were the two temporal test focus conditions. Similarly, even though the TCE was positively correlated with recall in all conditions, this correlation was greater for the temporal/temporal condition than either semantic test focus condition. Participants who most effectively followed their instructions for the List 2 test were also those who had the most recall success.

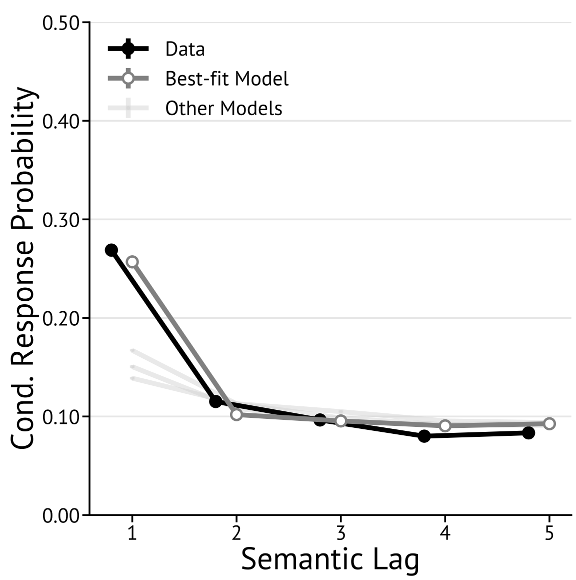

Figure 4.14 Behavioral data for the younger adults from PEERS Experiment 1 alongside simulated semantic-conditional response probabilities (semantic-CRPs) from model fits to semantic-CRPs only. Open points represent simulated data based on the best-fitting model of the Context Maintenance and Retrieval model. Additional model fits are plotted in light gray.

Finally, the computational implementation of retrieved context models that was fit to the data was able to simulate levels of overall recall and temporal contiguity well in all conditions but struggled to simulate data consistent with the observed level of semantic contiguity, especially in the semantic/semantic condition. This failure is somewhat surprising, as CMR has successfully fit to semantic contiguity in other datasets. However, the difference could be precisely due to a difference in datasets. The materials for this study were created in a similar way to that of the PEERS study, a dataset this version of CMR previously successfully fit (Morton & Polyn, 2016), but there is a key difference: directions during recall. In the PEERS study, all participants completed free recall whereas participants in Experiment 3 were instructed to organize their recalls according to their assigned goal. In particular, the semantic/semantic condition displayed a higher conditional probability of making a transition of semantic $lag = 1$ than participants in the PEERS study (PEERS behavioral data is presented in Figure 4.14 in black).

To test this explanation, I fit the model 5 times to behavioral data from PEERS Experiment 1. The simulated data for the best-fitting model is presented in Figure 4.14 alongside the behavioral data for the younger adult group in this experiment. The model was able to fit to the semantic contiguity in this dataset but not consistently across fits. Therefore, it is also possible that the model's difficulty is in fitting the particular measure of semantic contiguity used here, as a different measure was used by previous researchers fitting CMR to semantic contiguity (Morton & Polyn, 2016; Polyn et al., 2009a).

The results of Experiment 3 are also best explained by a combination of automatic mechanisms within retrieved context models and strategic control processes. Participants demonstrated the ability to flexibly change strategies from encoding to retrieval without any detriment to recall success or organization, indicating that both temporal and semantic associations are fully accessible at retrieval regardless of goals during encoding. The observed differences in temporal and semantic organization are therefore primarily a result of control processes at retrieval, enabling participants to choose different routes as they search their memories.

CHAPTER 5

HOW DO EXTERNAL GOALS AFFECT RECALL ORGANIZATION?

Experiment 4a

Although the TCE is extremely well-replicated, our understanding of how *influential* temporal information is in organizing memories is somewhat limited because most studies investigating temporal contiguity have used relatively artificial situations. In a typical free recall experiment examining the TCE, participants study a list of random, unrelated words and are then asked to recall those words in whatever order they come to mind. One advantage of a free recall task is that allowing participants to recall items *in any order* can offer insight into the basic organization of the memory system and may reflect automatic processes in a way that goal-directed recall (e.g., "recall the items in the order you saw them") does not. Conducting research on recall of simple stimuli, like lists of unrelated words, allows for a high degree of experimental control; the experimenter can minimize the effects of individual differences in participants' pre-existing associations between items and examine how temporal associations are used when the items are unrelated in any other way.

Yet, this paradigm also limits the generalizability of the TCE in two ways. First, in a free recall task, participants' goal is simply to recall as many words as possible. When recall is in the service of a more realistic goal that encourages a different kind of organization, it is unclear if participants will still use temporal information to guide their recalls. Second, it is possible that the TCE has been consistently observed in free recall not because temporal context information is automatically encoded and retrieved, but rather because temporal associations are the only associations available to participants (Hintzman, 2011). In a more realistic situation where items are related in many different ways, temporal organization may be less helpful, and the TCE may be eliminated.

Presence of an External Goal

As discussed in Chapter 4, recall organization can be greatly influenced by participants' intentional retrieval strategies, whether those strategies are spontaneously adopted by participants or are assigned by the experimenter. Such strategies might include recalling items in semantic versus temporal order or trying to group recalls by semantic category. Yet, outside of the lab retrieval is generally in service of some goal beyond simply remembering items in a certain order. The current set of task goals may determine how participants approach memory search. For example, when recalling the events of a recent vacation to a friend, it may be useful to recall the events in order for the purpose of telling a coherent story, even though temporal organization itself is not the goal. Temporal order may be irrelevant to other goals, such as recalling the best restaurants from that vacation.

Retrieved context models make the clear prediction of a TCE under any circumstance where new memories are formed. However, it is empirically less clear if temporal information influences recall when external goals emphasize the use of other kinds of associations. If the TCE is eliminated when other associations are emphasized by participants' goals, this would be a serious challenge to these models. These models should be able to simulate organization along multiple dimensions, such as membership in a specific category, if other kinds of associations are available.

Presence of Other Associative Dimensions

Another factor that has a strong influence on recall organization is the presence of non-temporal associations. The TCE is greatest in lists of random, unrelated words, where only temporal associations are available and useful for guiding recall. Does the TCE still occur when other associations are present and more useful, particularly if those associations are more relevant to the task goals? Supporting the retrieved context models assumption that temporal information is *always* encoded and retrieved, a robust TCE has been observed with recall of naturalistic stimuli, such as news stories, tours, and autobiographical memories (Diamond & Levine, 2020; Moreton & Ward, 2010; Pathman et al., 2023; Uitvlugt & Healey, 2019). However, memories formed nearby in time are often also associated in other ways (Buzsáki & Llinás, 2017; Diamond & Levine, 2020; Hintzman, 2016; Moreton & Ward, 2010; Rogerson et al., 2014; Silva et al., 2009; Uitvlugt & Healey, 2019). For example, animals seen close together in time during a trip to the zoo are also likely to be housed in nearby locations and be meaningfully related to one another (e.g., animals from the African Savannah may have been visited around the same time and in nearby locations; Pathman et al., 2023). Because the TCE is enhanced when other associations are correlated with time (Healey et al., 2019), the TCE observed with most naturalistic stimuli is likely influenced by other kinds of associations, and therefore it cannot be used as a measure of pure temporal organization.

When related items are not presented nearby in time, the TCE is reduced. For example, when participants study a list that contains multiple semantic clusters (groups of 3-4 words that are semantically related to each other, but not to other items in the list) and semantically related words are not presented nearby in time, the TCE is reduced compared to recall of random word lists, and participants instead display a strong semantic contiguity effect (Healey & Uitvlugt, 2019; Hong et al., 2022; Polyn et al., 2011). Yet, even in lists where items are semantically related, the relationships between the items are often relatively simplistic (e.g., category membership). The patterns of recall organization observed when related items are not experienced in temporal order may not hold for more complex relationships between items. This may be especially true if temporal information does not hold a central and foundational role in the episodic memory system but is merely a strategy used in free recall when no other resources are available, as some have suggested (Hintzman, 2016). If temporal organization is primarily situational, then when a richer network of associations is available, participants should be drawn to other associations and abandon their simplistic temporal strategies. On the other hand, if temporal context forms an important part of all episodic memories, then memories will still be organized based on temporal order, even in the presence of other associations and external goals.

In the present experiment, participants studied lists of food items with one of three goals: remember the items for a later shopping trip, remember the items for cooking several dishes later, or remember as many items as possible. All participants knew their memory would be tested; the key difference between conditions was in their external goal. Each food item was paired with one of three shopping locations (farmer's market, supermarket, or specialty store) and one of three dishes (appetizer, main, or side). Therefore, participants could organize their recalls based on location, dish, or temporal order. The goal of this experiment was to test if there is still temporal contiguity in participants' recalls when other associations are available and task-relevant and to test one way in which task goals can be implemented in retrieved context models.

Methods

Materials

Each participant studied 2 lists of 18 item-location-dish triplets. The item for each triplet was randomly selected without replacement for each participant from a list of 100 possible ingredients, a subset of the larger word pool developed for PEERS (see Healey et al., 2019). The pool of possible ingredients were food words that could be plausibly used as an ingredient to cook a multi-ingredient dish for a meal (e.g., CHICKEN, APPLE, ZUCCHINI). All of the items in the word pool were compared to ensure they did not represent a general category that would also include other items in the pool. For example, PORK and CHICKEN were both possible ingredients, but MEAT was not included in the word pool because MEAT would include both PORK and CHICKEN. Already-prepared food items, such as PASTA or DOUGHNUT, were also excluded from the word pool.

During study, each ingredient was presented together with both a location where that ingredient could be purchased (farmer's market, supermarket, or specialty store) and the type of dish the ingredient would be used to prepare (appetizer, main, or side). A brief description of each of these locations and dishes was provided to participants during the initial instructions. The ingredient was always presented on top in a larger font, and the location and dish were presented below the ingredient. The order of the location and dish (top/bottom) was counterbalanced across participants. The triplets were arranged in random order, and each location/dish combination was presented exactly twice for each participant.

Design

Participants were randomly assigned to one of three conditions in a between-subjects design. In the location-focus condition, participants were instructed prior to encoding that their goal was to memorize the list of ingredients for a shopping trip where they would travel to each of the three kinds of stores (farmer's market, supermarket, specialty store). In the dish-focus condition, participants were instructed prior to encoding that their goal was to memorize the list of ingredients for when they would cook the three dishes at home (appetizer, main, or side). In the control free recall condition, participants were informed that each item would be paired with a location where it could be purchased and a dish it would be used to make, but they were simply told their goal was to recall as many items as possible in whatever order they came to mind. None of the participants were explicitly instructed on how to group their recalls; the location-focus and dish-focus conditions were just given a goal which could encourage participants to organize their recalls based on the location or dish paired with each ingredient. All participants were instructed to try to remember as many words as possible.

Participants

A sample size of 140 per condition would provide 95% $1 - \beta$ power to detect a TCE of the size found by Polyn et al. (2011) in lists with a similar semantic structure. Because in the present experiment participants studied only 2 lists, compared to 15 lists per condition in Polyn et al. (2011), and the TCE tends to increase with practice (Healey et al., 2019), I aimed to collect data from at least 600 participants (200 per condition) to allow detection of an even smaller effect. In total, 612 Michigan State University undergraduate students completed this experiment online on their personal computers.

Because the key manipulation of this experiment requires participants to attend to their assigned goal, all participants were asked on a post-experiment questionnaire if there was any reason their data should not be included. Thirty-five participants who responded "YES" to this question were excluded. Most participants who responded "YES" provided a reason why their response should be excluded. For example, some reported there was a large interruption while they were completing the experiment, they did not understand the instructions, or they cheated on the memory test by writing some of the words down as they studied them. In addition, any participant who recalled more than 36 items (i.e., more than twice the length of the list) was excluded from analysis because it was unlikely they were following the task instructions. In total, three participants were excluded for recalling more than 36 items on at least one of the lists. Of the remaining 574 included participants, 403 (70.2%) reported their gender as female, and the mean age was 20.0 years (SD = 1.9).

Procedure

At the beginning of the experiment, participants were informed that they would be studying 2 lists of food items, and their goal would be to recall as many of the items as possible on a later test. They were asked to "Imagine you are planning a shopping trip to buy ingredients for three different dishes: an appetizer, a main dish, and a side dish." Participants were informed that each ingredient could be purchased at *one* of three locations and would be used in making *one* dish.

The next set of instructions differed by condition. Participants in the location-focus condition were told "Your goal is to remember as many of the ingredients as you can FOR YOUR SHOPPING TRIP TO THE THREE LOCATIONS." They were instructed that on the memory test following each list they should "Try to imagine you are listing the ingredients as you shop. As you travel from store to store, which items do you remember?" Participants in the dish-focus condition were told "Your goal is to remember as many of the ingredients as you can for WHEN YOU COOK THE DISHES." They were instructed that on the memory test following each list to "Try to imagine you are listing the ingredients after you have brought them home. As you are cooking the dishes, which items do you remember?" In the free recall condition, participants were simply told "Your goal is to remember as many of the ingredients from the list as you can." They were instructed that on the memory test following each list they should "Type the words in whatever order they come to mind." A reminder of each participant's goal (shopping, cooking, or free recall) also appeared at the beginning of List 2 before study.

Each participant studied 2 lists of 18 item-location-dish triplets. Each triplet was presented for 5 s with a jittered inter-stimulus interval of 400-600 ms. The study period was followed by a short delay where participants read a reminder of their instructions before proceeding to a free recall screen where they could type in any words they could remember individually into a provided text box. After typing each word, they pressed ENTER, and the word disappeared, leaving a blank text box for the next recall. Participants had 2 min for recall.

After recalling the second list, participants completed a short post-experiment survey where they provided demographic information and answered questions about their strategies during each list. They were also asked to report if there was any reason to exclude their data.

Results

To fully examine the effects of external goals in a task where items are related along multiple dimensions, I examined measures of both recall performance and recall organization. All analyses reported here are averaged over both lists.[1]

Recall Performance

Recall performance was measured using recall probabilities, serial position curves, and probability of first recall curves, presented in Figure 5.1. On average, participants in all conditions had high levels of recall success compared to previous online studies with the same population where recall probabilities are often below 0.4 (e.g., Chapter 2; Healey & Uitvlugt, 2019). Recall was likely higher in the present study because of the categorized structure of the list (e.g., Healey & Uitvlugt, 2019). Recall probabilities, displayed in Figure 5.1A, were greater than 0.5 in the free recall ($M = 0.52$, $SD = 0.19$), location-focus ($M = 0.54$, $SD = 0.17$), and dish-focus ($M = 0.54$, $SD = 0.16$) conditions. However, there was no effect of task goal on overall recall, $F(2, 571) = 0.93$, $p = .397$.

Although there were no differences in overall recall, it is also important to consider other, more detailed measures of recall which may reveal differences between conditions. Probability of first recall curves (Figure 5.1B), represent the probability of initiating recall from each serial position in the list. Participants displayed a strong tendency to initiate recall from the beginning of the list in all conditions. The serial position curves for all conditions (Figure 5.1C) also display a clear pattern of strong primacy with minimal recency (for similar patterns, see Bjork & Whitten, 1974;

[1]Recall and temporal contiguity were slightly higher for List 2 for all three conditions, consistent with previous work finding that the TCE increases with task experience (Healey et al., 2019). Semantic contiguity, clustering by location, and clustering by dish were similar in both lists across conditions. More importantly, the relationships between the conditions were no different for List 1 and List 2. Therefore, all analyses presented here are averaged across both lists.

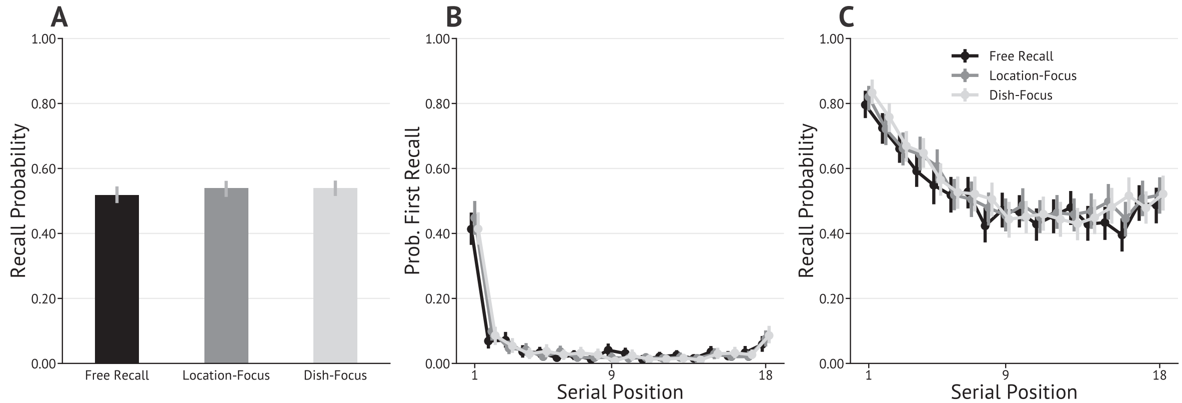

Figure 5.1 Recall performance by condition, measured with (A) recall probabilities, (B) probability of first recall curves, and (C) serial position curves. Error bars represent bootstrapped 95% confidence intervals.

Neath, 1993; Unsworth, 2008). Assigning a goal did not affect either overall or recall of particular items.

Recall Organization

Temporal Contiguity. The main aim of this experiment is to determine if the TCE is eliminated when other kinds of associations are available and participants are provided with a realistic goal that encourages the use of those other associations to guide recall. I examined temporal contiguity using both chance-adjusted temporal factor (TF) scores and temporal bias scores (for detailed descriptions of these analyses, see Chapters 3 and 2, respectively). The TCE was substantially above chance in all conditions (indicated by chance-adjusted TF scores greater than zero in Figure 5.2A). The chance-adjusted TF scores clearly demonstrate that temporal order information did guide recall even in the presence of other task-relevant associations[2]. In fact, chance-adjusted TF scores were similar across conditions (free recall: $M = 1.15$, $SD = 1.28$; location-focus: $M = 1.16$, $SD = 1.20$; dish-focus: $M = 1.10$, $SD = 1.17$), and there was no effect of task goal, $F(2, 565) = 0.11$, $p = .900$. Temporal bias scores, which provide a more detailed measure of temporal contiguity, are presented

[2]Calculating chance-adjusted TF scores requires participants to recall at least 2 list items in a given list so the chance distribution will have a standard deviation of more than zero. If at least 2 items are recalled, because positive and negative lags are treated as different values in this calculation, then at least 2 different TF scores are possible for each permutation. Chance-adjusted TF scores could not be calculated for the 6 participants who recalled fewer than 2 list items on both lists, so they were excluded from this analysis.

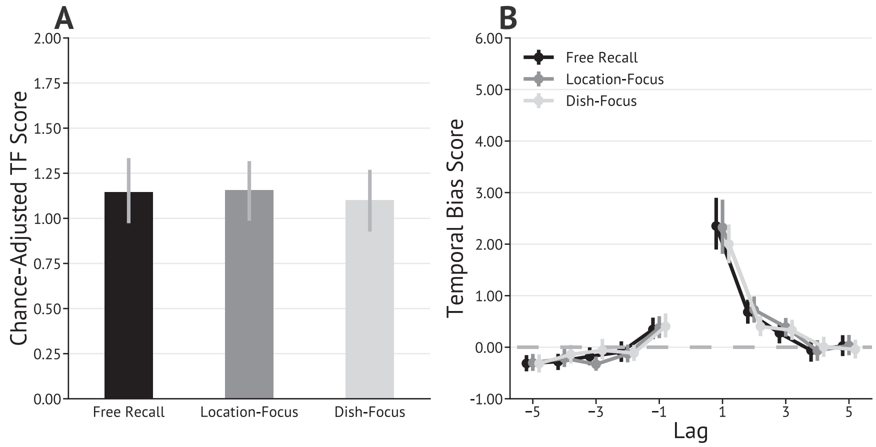

Figure 5.2 Temporal contiguity by condition, measured with (A) chance-adjusted temporal factor (TF) scores, and (B) temporal bias scores. For TF scores, chance was determined by permuting the order of recalls 1,000 times. Chance-adjusted TF scores were calculated for each participant by subtracting the average of the chance distribution from the actual TF score for that participant and dividing by the standard deviation of the chance distribution. Temporal bias scores for each lag were calculated for each subject by comparing the number of times a transition of that lag was actually made to the number of times it would be expected to occur by chance. Chance was calculated by permuting the order of recalls for each list 1,000 times and counting on average how many times each lag occurred for each permutation. The dotted line for the temporal bias scores indicates a score of zero (no bias). Error bars represent bootstrapped 95% confidence intervals.

in Figure 5.2B. Participants displayed a clear TCE with a strong bias for recalling items in forward order, as is typically observed in free recall of random lists (Healey et al., 2019). This pattern was consistent across all three conditions, further demonstrating that the TCE was unaffected by task goals.

Participants in all conditions displayed temporal organization, even when other kinds of associations were present and task-relevant. But did goals affect recall organization along other dimensions? I also examined recall organization based on semantic similarity, location category, and dish category.

Semantic Contiguity. Semantic contiguity was measured using chance-adjusted semantic factor (SF) scores and semantic lag-conditional response probabilities (semantic-CRPs; for descriptions see Chapters 3 and 4, respectively). Chance-adjusted SF scores, presented in Figure 5.3A) were low but above zero in all conditions (free recall: $M = 0.33$, $SD = 0.71$; location-focus:

$M = 0.22$, $SD = 0.74$; dish-focus: $M = 0.36$, $SD = 0.75$).[3] That is, participants in all three conditions displayed some level of semantic organization. Because the lists were all composed of food items, semantically associated words were present in each list, making it possible for participants to organize their recalls based on these associations. However, semantic contiguity did not differ by condition. There was no effect of task goals on chance-adjusted SF scores, $F(2, 564) = 1.99$, $p = .138$. The semantic-CRPs, which provide a more detailed measure of semantic organization, also did not detect any differences between conditions (Figure 5.3B). Participants in all three conditions were slightly more likely to make transitions to the next most semantically similar item (semantic *lag* = 1) than to less semantically similar items. Task goals also did not affect the degree of semantic organization.

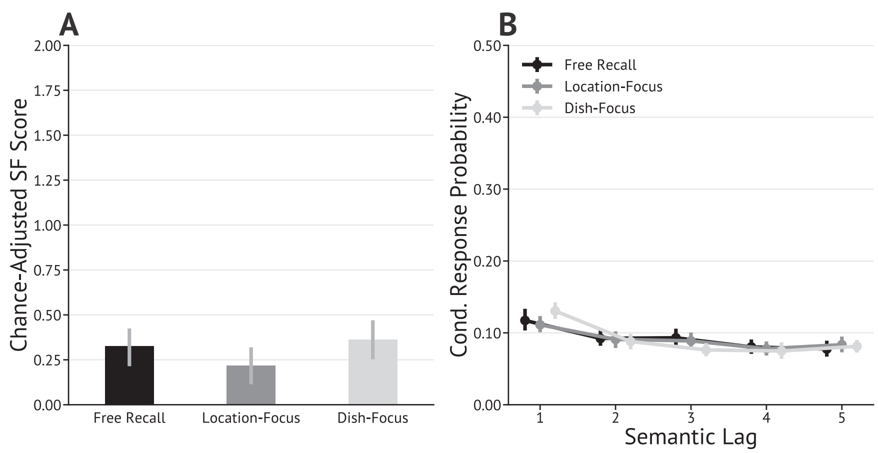

Figure 5.3 Semantic contiguity by condition, measured with (A) chance-adjusted semantic factor (SF) scores, and (B) semantic lag-conditional probabilities (semantic-CRPs). For SF scores, chance was determined by permuting the order of recalls 1,000 times. Chance-adjusted SF scores were calculated for each participant by subtracting the average of the chance distribution from the actual SF score for that participant and then dividing by the standard deviation of the chance distribution. Error bars represent bootstrapped 95% confidence intervals.

Category Clustering. To determine the degree to which participants organized their recalls based on location and dish categories, I examined multiple measures of category clustering. For

[3] Calculating chance-adjusted SF scores requires participants to recall more than 2 list items for a given list so the chance distribution will have a standard deviation of more than zero. If 2 or fewer items are recalled, because semantic lags are only positive (in contrast to temporal lags, which can be both positive and negative), only one SF score is possible, no matter how the values are permuted. Chance-adjusted SF scores could not be calculated for the 7 participants who recalled 2 or fewer list items on both lists, so they were excluded from this analysis.

the purpose of these analyses, recalling two items in a row that were associated with the same location (for location clustering measures) or dish (for dish clustering measures) was considered a within-category transition. There were three possible categories for location (farmer's market, supermarket, specialty store) and three possible categories for dish (appetizer, main, side). Any transitions to or from an intrusion were excluded from these analyses for consistency with the temporal and semantic analyses. Each analysis was conducted separately for location clustering and dish clustering.

Adjusted ratio of clustering (ARC) scores are commonly used as a measure of category clustering (Roenker et al., 1971). ARC scores compare the number of transitions between items in the same category in participants' actual recalls to the number of within-category transitions that would be expected, given the number of items and number of categories recalled. Scores above zero indicate clustering above chance, with an upper bound of 1.0, and scores below zero indicate clustering below chance. An advantage of this measure is that it accounts for chance by not only considering the actual items recalled but also the number of categories that were actually recalled and the number of category repetitions that are possible given the participant's actual recalls.

To provide another measure of category clustering that is more comparable to the measures of temporal and semantic contiguity reported here and in the previous chapters, I also calculated chance-adjusted category clustering scores. For chance-adjusted category clustering scores, I calculated for each participant and for each list the number of within-category transitions to get the actual number of transitions. The number of actual within-category transitions was divided by the number of times a within-category transition was possible. A within-category transition was considered possible unless all of the items in the just-recalled category had already been recalled. For example, if the participant had already recalled all of the ingredients for the main dish, then a transition to another main dish item would not be possible. The number of times a within-category transition was made was divided by the number of times a within-category transition was possible for each participant and each list to calculate a category clustering score. Chance was calculated in a similar way as for chance-adjusted temporal and semantic factor scores; for each participant and for each list, their recalls were permuted 1,000 times, and chance was calculated as the probability of making a within-category transition across all 1,000 permutations (i.e., if the same items were recalled in random order). The final chance-adjusted clustering score was calculated as the actual clustering score minus the chance clustering score, divided by the standard deviation of the chance clustering score.

Both of these measures consider the number of within-category clusters relative to the categories of the items that were actually recalled and therefore would be expected to lead to similar conclusions. ARC scores are commonly used in the literature to measure clustering by category; I used them here to facilitate comparisons between the present experiment's results and previous work. The chance-adjusted clustering scores are a novel measure of category clustering that is more comparable to the chance-adjusted SF and TF scores used to measure semantic and temporal organization in this dissertation. For this reason, both ARC scores and chance-adjusted clustering scores are reported. ARC scores for location and dish clustering are presented in Figures 5.4B and 5.5B, and chance-adjusted clustering scores for location and dish clustering are presented in Figures 5.4A and 5.5A.

There was no effect of condition on location clustering, regardless of whether clustering was measured with ARC scores, $F(2, 563) = 0.78$, $p = .461$, or chance-adjusted clustering scores, $F(2, 562) = 1.32$, $p = .268$. ARC scores for clustering were no different from chance, as indicated by the bootstrapped 95% confidence intervals overlapping zero, in all three conditions (free recall: $M = 0.04$, $SD = 0.29$; location-focus: $M = 0.02$, $SD = 0.28$; dish-focus: $M = 0.02$, $SD = 0.25$). Similarly, chance-adjusted clustering scores overlapped with chance for free recall ($M = 0.08$, $SD = 0.73$), location-focus ($M = -0.03$, $SD = 0.74$), and dish-focus ($M = -0.02$, $SD = 0.73$).[4]

Dish clustering was also unaffected by task goal. There was no effect of condition on clustering

[4] As with chance-adjusted TF and SF scores, ARC and chance-adjusted clustering scores could not be calculated for some participants. Chance-adjusted clustering scores could only be calculated when the standard deviation of the chance distribution was $\neq 0$. Therefore, 9 participants were excluded from the chance-adjusted clustering score analyses. ARC scores could not be calculated for participants whose maximum number of within-category transitions was equal to the expected (chance) number of within-category transitions which can occur if 1) a participant recalls 2 or fewer items in a row or 2) items from only 1 category are recalled. In total, location ARC scores could not be calculated for 8 participants, and dish ARC scores could not be calculated for 9 participants.

using either ARC scores, $F(2, 562) = 0.77$, $p = .463$, or chance-adjusted clustering scores, $F(2, 562) = 1.84$, $p = .160$. ARC scores were no different from chance for the free recall condition ($M = 0.04$, $SD = 0.32$), location-focus condition ($M = 0.04$, $SD = 0.26$), and dish-focus condition ($M = 0.01$, $SD = 0.25$). Chance-adjusted dish clustering scores were also at chance in all three conditions (free recall: $M = 0.06$, $SD = .77$; location-focus: $M = 0.07$, $SD = 0.78$; dish-focus: $M = -0.06$, $SD = 0.65$).

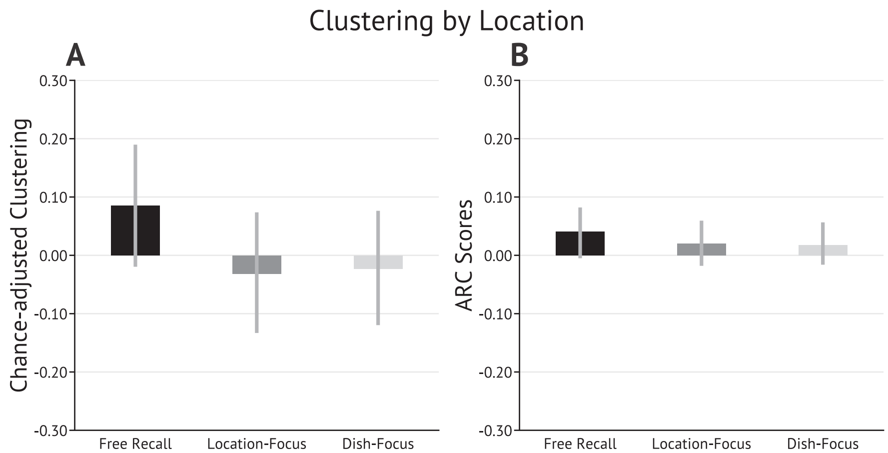

Figure 5.4 Clustering by location category for all conditions, measured with A) chance-adjusted clustering scores and B) Adjusted Ratio of Clustering (ARC) scores, two measures of the probability of recalling two items in a row associated with the same category (e.g., two items associated with SUPERMARKET). Chance-adjusted clustering scores were calculated by counting the total number of within-category transitions and dividing by the number of times a within-category transition was possible. This clustering score was then compared to chance, which was determined by permuting the order of recalls 1,000 times, calculating clustering scores for each permutation to get a measure of the degree of clustering that would be expected if the same items were recalled in random order. Chance-adjusted clustering scores were calculated for each participant by subtracting the average of the chance distribution from the actual clustering score for that participant and then dividing by the standard deviation of the chance distribution. For both ARC scores and chance-adjusted clustering scores, a score of zero represents chance. Error bars represent bootstrapped 95% confidence intervals.

Individual Differences

On average, participants did not organize their recalls either by associated location or dish. However, some participants may have organized their memory search more effectively than others. For example when semantic associations are minimized, temporal information is positively cor-

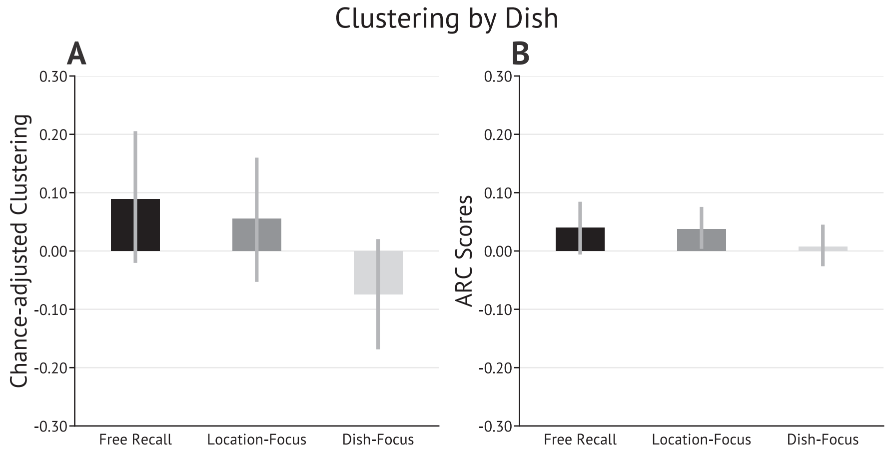

Figure 5.5 Clustering by dish category for all conditions, measured with A) chance-adjusted clustering scores and B) Adjusted Ratio of Clustering (ARC) scores, two measures of the probability of recalling two items in a row associated with the same category (e.g., two items associated with APPETIZER). Chance-adjusted clustering scores were calculated by counting the total number of within-category transitions and dividing by the number of times a within-category transition was possible. This clustering score was then compared to chance, which was determined by permuting the order of recalls 1,000 times, calculating clustering scores for each permutation to get a measure of the degree of clustering that would be expected if the same items were recalled in random order. Chance-adjusted clustering scores were calculated for each participant by subtracting the average of the chance distribution from the actual clustering score for that participant and then dividing by the standard deviation of the chance distribution. For both ARC scores and chance-adjusted clustering scores, a score of zero represents chance. Error bars represent bootstrapped 95% confidence intervals.

related with recall performance; however, the correlation between temporal contiguity and recall performance is reduced or even eliminated when strong semantic associations are present (Healey & Uitvlugt, 2019; Hong et al., 2023; Sederberg et al., 2010). It is unclear if the same pattern would occur with materials like those in this experiment. To test if recall organization was consistent across individuals or if high-performing participants engaged in different kinds of organization than low-performers, I examined correlations between recall success and each measure of recall organization.

The relationships between each measure of organization and recall success are presented in Figure 5.6 for the free recall condition, Figure 5.7 for the location-focus condition, and Figure 5.8 for the dish-focus condition. Split-half reliabilities are presented in Table 5.1 for each measure. Although reliabilities were fairly low for all measures except recall probability, there are two points

Split-half Reliability for Individual Difference Variables

Table 5.1 Split-half reliability for recall probability, chance-adjusted temporal factor (TF) scores, chance-adjusted semantic factor (SF) scores, Adjusted Ratio of Clustering (ARC) scores (for both location and dish category clustering), and chance-adjusted clustering scores (for both location and dish category clustering) are presented here. For each condition, split-half reliability was calculated following the methodology of Sederberg et al. (2010). For each participant with 2 valid lists (where at least 2 list items were recalled), I randomly divided the participant's lists into two sets. I calculated recall probability, chance-adjusted factor scores, ARC scores, and chance-adjusted clustering scores for each set and correlated the scores for set 1 with scores for set 2, correcting with the Spearman-Brown prediction formula ($2\rho/[1+\rho]$).

Measure	Free Recall	Location-Focus	Dish-Focus
Recall Prob.	0.730	0.594	0.641
Chance-adjusted TF Score	0.496	0.255	0.353
Chance-adjusted SF Score	0.049	-0.088	0.047
ARC Score (Location)	0.013	0.268	-0.024
ARC Score (Dish)	0.088	-0.044	-0.237
Chance-adjusted Clustering (Location)	0.061	-0.049	0.134
Chance-adjusted Clustering (Dish)	-0.020	0.064	-0.345

worth noting with regards to individual differences.

First, only temporal organization predicted recall success. Chance-adjusted TF scores (top left of Figures 5.6, 5.7, and 5.8) were positively correlated with recall in all three conditions. As has been found in previous work with simpler stimuli, memory for items and memory for their order were linked together. Recall success was not predicted by chance-adjusted SF scores. However, SF scores were quite unreliable (Table 5.1), making these correlations difficult to interpret.

Second, clustering scores were not higher for high-performing participants. Although reliabilities for the clustering measures were low, scatter plots of ARC scores (second row of Figures 5.6, 5.7, and 5.8) and chance-adjusted clustering scores (third row of Figures 5.6, 5.7, and 5.8) display that clustering by category was near chance even for participants who recalled a high proportion of list items. One pattern did emerge in examining individual differences in recall organization and recall success. Greater variability was present in ARC scores for low-performing participants. That is, low-performing participants were more likely to have extremely high or extremely low ARC scores, while high-performing participants were more likely to have ARC scores near zero. It

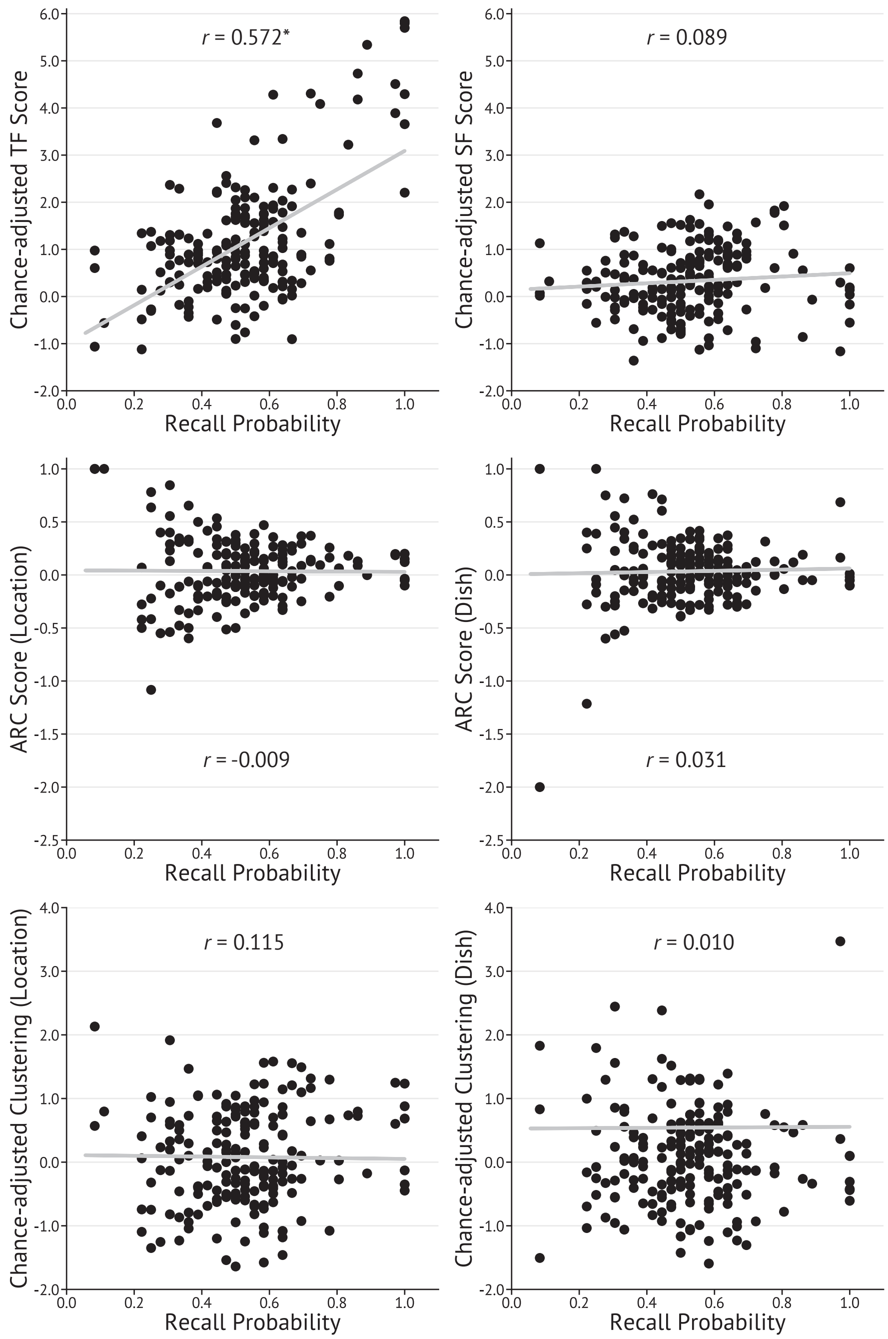

Figure 5.6 Correlations for the free recall condition between recall probability and each measure of recall organization: chance-adjusted temporal factor (TF) scores, chance-adjusted semantic factor (SF) scores, Adjusted Ratio of Clustering (ARC) scores for both dish and location category clustering, and chance-adjusted clustering scores for both location and dish category clustering. Pearson's r correlation coefficients are presented for each measure with the line of best fit plotted in gray. Correlations marked with a * are significant with $p < .0001$.

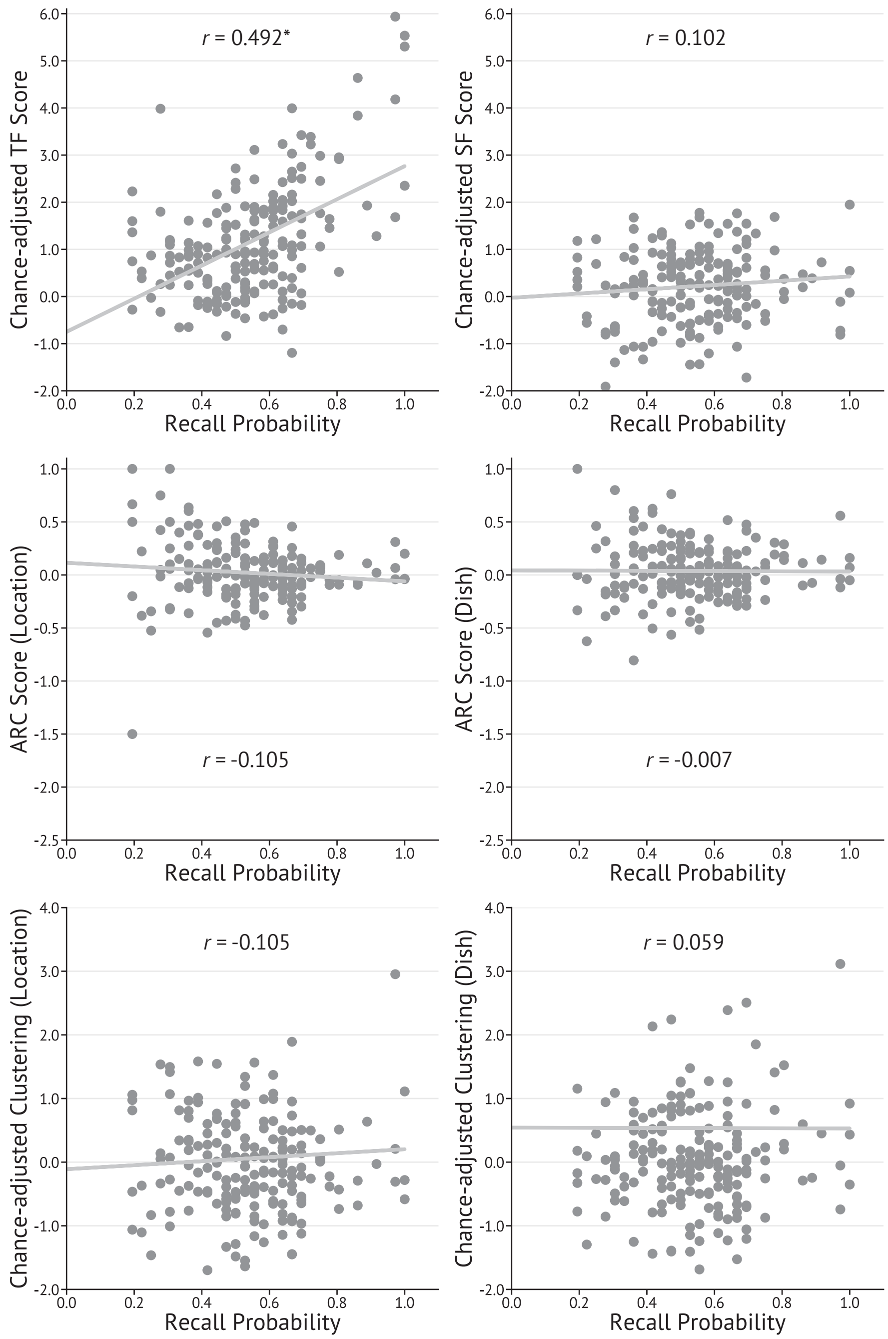

Figure 5.7 Correlations for the location-focus condition between recall probability and each measure of recall organization: chance-adjusted temporal factor (TF) scores, chance-adjusted semantic factor (SF) scores, Adjusted Ratio of Clustering (ARC) scores for both dish and location category clustering, and chance-adjusted clustering scores for both location and dish category clustering. Pearson's r correlation coefficients are presented for each measure with the line of best fit plotted in gray. Correlations marked with a * are significant with $p < .0001$.

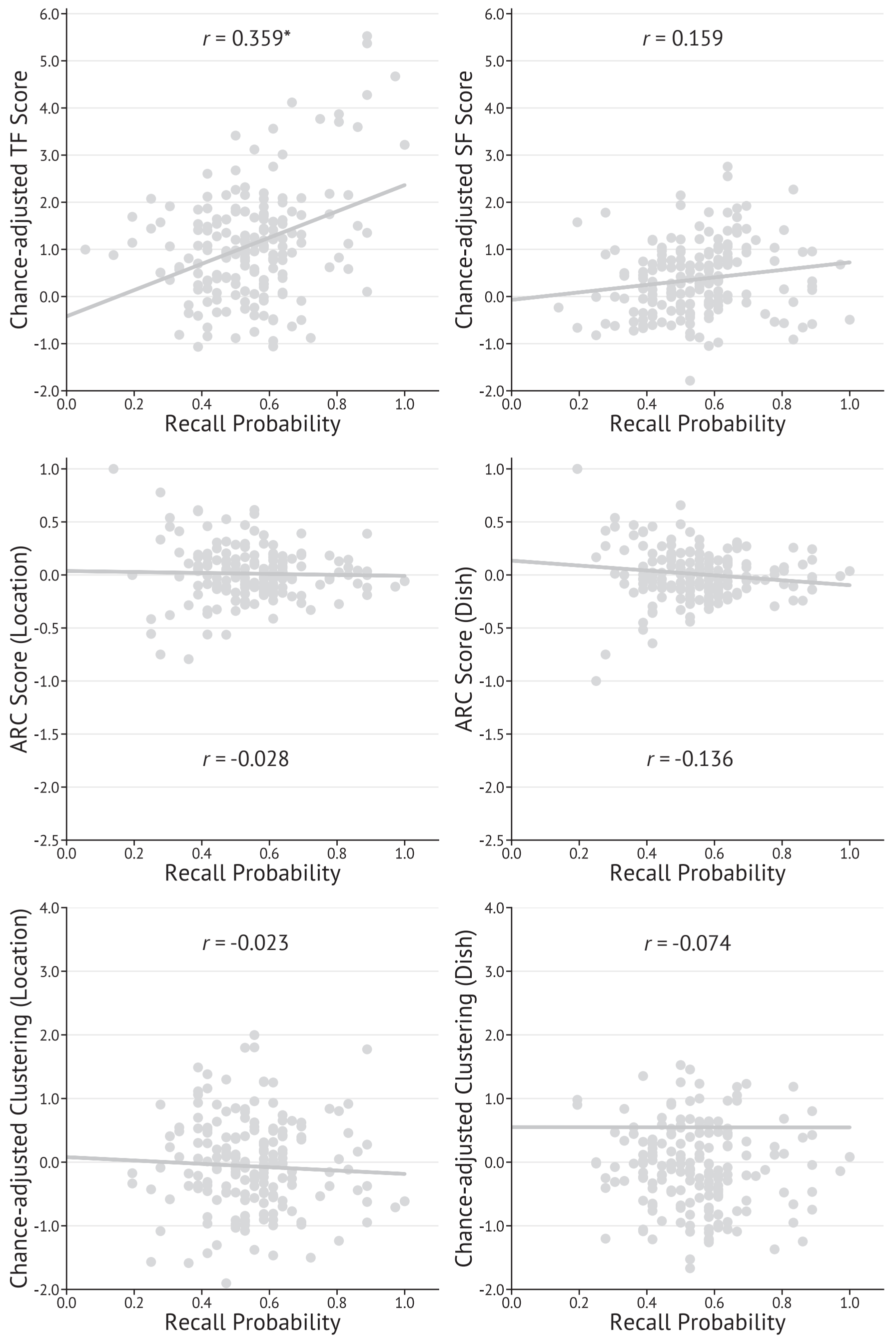

Figure 5.8 Correlations for the dish-focus condition between recall probability and each measure of recall organization: chance-adjusted temporal factor (TF) scores, chance-adjusted semantic factor (SF) scores, Adjusted Ratio of Clustering (ARC) scores for both dish and location category clustering, and chance-adjusted clustering scores for both location and dish category clustering. Pearson's r correlation coefficients are presented for each measure with the line of best fit plotted in gray. Correlations marked with a * are significant with $p < .0001$.

is possible that organizing recalls by location and dish was not an effective strategy, and therefore those participants who recalled the most words were those who ignored these associations. Another possibility is that that ARC scores are simply a poor psychometric measure. Regardless of the cause of these differences, some participants did engage in clustering by list or dish category, but on average, both kinds of category clustering were near chance.

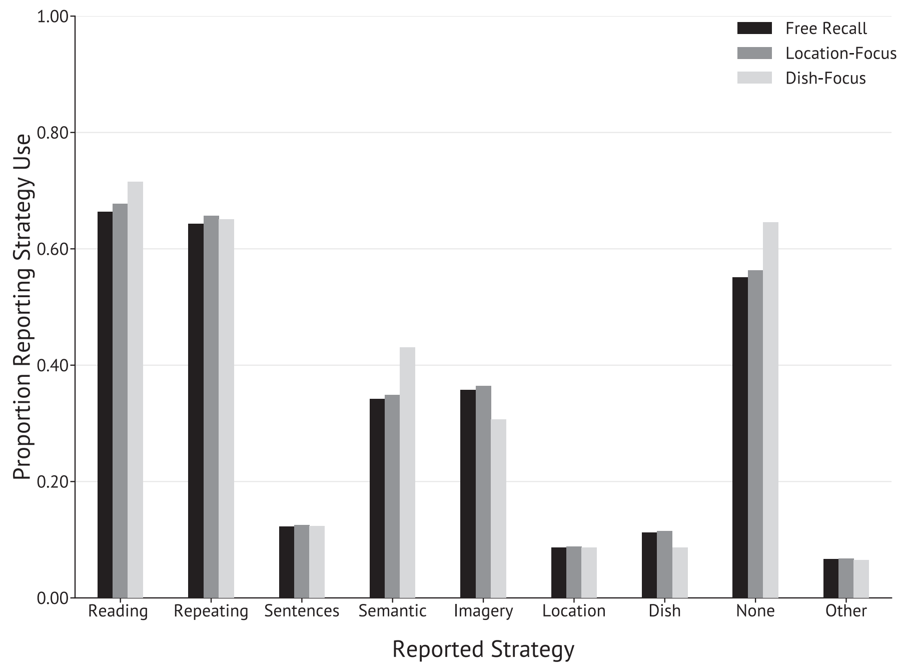

Figure 5.9 Encoding strategies reported by participants in Experiment 4a. The proportion of participants reporting using each strategy is plotted for each condition. Because participants were allowed to select multiple answers, the sums of the proportions for each condition are greater than 1.0.

Self-Reported Strategy Use

The lack of clustering by location and dish, although surprising, is consistent with the strategies participants reported on the post-experiment survey. On this survey, participants were asked to indicate the kinds of strategies they used during encoding and retrieval throughout the experiment. Specifically, participants were asked, “When you were trying to MEMORIZE the words, which strategies (if any) did you use? If you used more than one (e.g., you tried different strategies at different points), check all that apply.” Participants were allowed to select multiple strategies in

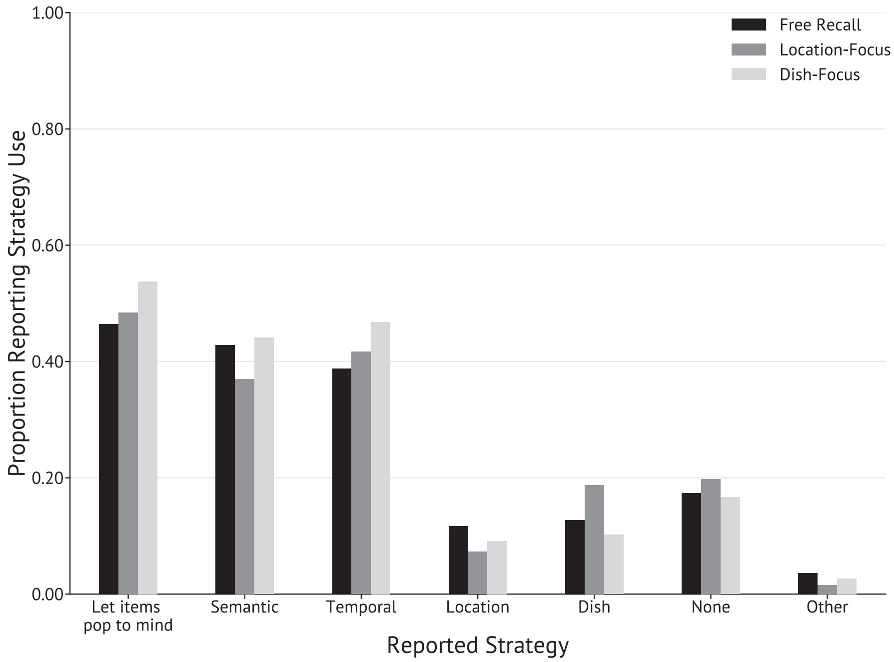

Figure 5.10 Recall strategies reported by participants in Experiment 4a. The proportion of participants reporting using each strategy is plotted for each condition. Because participants were allowed to select multiple answers, the sums of the proportions for each condition are greater than 1.0.

response to this question. The proportion of participants indicating that they used each encoding strategy is presented in Figure 5.9. A majority of participants reported using shallow strategies, like, "Reading each word as it appeared" and "Repeating the words as much as possible." However, less than 12% of participants in any condition reported focusing on the location or dish associations during encoding.

Participants were also asked, "When you were trying to RECALL the words, which strategies (if any) did you use to control the order in which words came to mind? If you used more than one (e.g., you tried different strategies at different points), check all that apply:" As displayed in Figure 5.10, very few participants reported trying to recall items based on their associated location or dish. Even in the location-focus condition, only 7.3% of participants reported using a location-based retrieval strategy, and in the dish-focus condition, only 10.2% of participants reported using a dish-based retrieval strategy. These responses are consistent with both the self-reported encoding strategy use and the chance-level category clustering observed in the ARC scores and chance-adjusted clustering scores.

Computational Modeling

The above analysis of the behavioral data provides a qualitative test of retrieved context models. The finding that the TCE was not eliminated in the presence of other associations and external goals was consistent with retrieved context models' predictions. However, another aim of this study is to take the first steps toward developing a computational model that is not only able to explain the presence of a TCE but also how external goals influence recall organization. Even though there was no effect of goals in this experiment, a comprehensive model of memory should be able to make predictions about the kinds of recall organization that would occur if task goals did determine recall organization.

To include task goals as a part of retrieved context models, I created a modified version of the Post-Encoding Pre-Production Reinstatement Model (PEPPR). For a description of the parameters in the version of PEPPR implemented here with goal representations, see Table 5.2. PEPPR, introduced by Healey and Wahlheim (2023), is a version of retrieved context models that includes representations of a label given to each list on the item layer (e.g., "LIST 1" and "LIST 2"). These representations allow the model to simulate recall dynamics resulting from goal-directed recall, such as when participants are directed to recall items only from List 2.

Other kinds of goals could also be represented in a similar way. An example encoding period for a modified version of PEPPR with goals is presented in Figure 5.11. Task goals (LOCATION FOCUS in the example in Figure 5.11) are represented as a node on the feature layer and context layer. The relevant goal node is activated when participants are assigned their goal and again when participants are reminded of their goal prior to retrieval. The rate of context drift when the goal is initially encoded and when participants are reminded of their goal prior to the retrieval period are controlled by the model parameters $\beta_{PEPPR_{encoding}}$ and $\beta_{PEPPR_{retrieval}}$ respectively (see Table 5.2).

To allow task goals to influence the retrieval of task-relevant information, I also added representations of the location and dish associated with each item as nodes on the feature layer. Thus, when the model encodes each item, it encodes the to-be-remembered ingredient, followed by its associated location, and then its associated dish. All of these are represented on the feature layer, form an association with the current state of context, and reinstate their own pre-experimental associations, causing context to drift. For example, in Figure 5.11, when the first triplet, ONION-SUPERMARKET-SIDE, is studied, ONION is activated on the feature layer. ONION then activates its associated context, represented with the image of an onion on the temporal context layer, and context drifts such that the goal (LOCATION FOCUS) is less active in the context layer than it was previously. The encoding process occurs again for the location (SUPERMARKET) and dish (SIDE) in the same way as for ONION. When the next triplet appears, encoding of the ingredient, location, and dish proceeds in the same way.

Because semantic organization was not a main focus of this modeling endeavor, semantic associations between the items were not considered. However, in PEPPR with goals each location is pre-experimentally associated with the LOCATION goal, and each dish is pre-experimentally associated with the DISH goal. The strength of these associations is controlled by the model parameter ϵ (see Table 5.2). The pre-experimental associations between the goals and their relevant features may allow for a higher likelihood that task-relevant features will be retrieved, as described below.

At recall, if the model was assigned a specific goal (location-focus or dish-focus), the representation of that goal is re-activated. The context associated with each goal is then reinstated, which provides a good cue for items studied nearby in time to the original goal assignment because of new associations formed during study *and* to task-relevant features because of their pre-experimental associations. The retrieval process then continues as described in Appendix A. If a to-be-remembered ingredient is retrieved, the model outputs the ingredient and continues to the next recall. If a location or dish is retrieved by the model, its associated context is also reinstated, incorporated into the current state of context, and used as a cue for the next recall. However, the recalled location or dish is not recorded as an output. This process continues until a to-be-remembered item is retrieved or recall fails, in which case recall ends for that list.

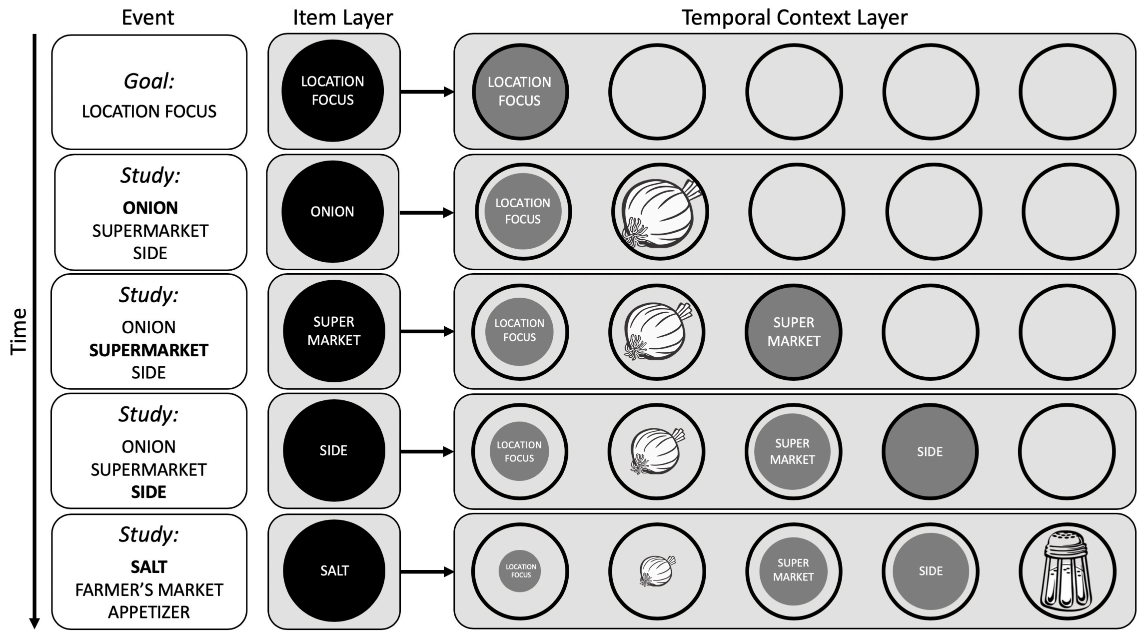

Figure 5.11 A visual example of the encoding period for a version of the Post-Encoding Pre-Production Reinstatement Model (PEPPR) with goal representations. The identities of items are represented on the item layer, with one node for each item. Each item has a corresponding node on the context layer which represents that item's contextual associates. First, the goal is assigned (*location focus*). The node for the *location focus* goal is activated on the item layer, and it activates its associated context node on the temporal context layer. Then, when the first triplet is studied, the model encodes the ingredient first, followed by the location and the dish. The node corresponding to the studied ingredient *onion* becomes active on the item layer, completely replacing the *location focus* activation on the item layer. *Onion* then activates its associated node on the context layer but does not entirely replace the *location focus* context. Instead, the *location focus* context fades so that the current mental context is a blend of the goal (*location focus*) and the just-studied ingredient (*onion*) in which the more recent event is more strongly represented. This process repeats as the model processes the associated location, *supermarket*: its respective node on the item layer becomes active, and its context representation is activated on the context layer, blending with other elements on the context layer to create a new *location focus-onion-supermarket* context. The same process repeats again for encoding of the associated dish (*side*). This sequence repeats for each of the triplets in the list.

Modeling Methods

To test this novel version of PEPPR's predictions for how task goals should affect recall organization, I fit the model to the overall recall probabilities, temporal bias scores, and chance-adjusted clustering scores for the free recall condition. For this condition, the "free recall" goal node was activated, which was not pre-experimentally associated with any of the locations or

Parameters for the Post-Encoding Pre-Production Reinstatement Model (PEPPR) to the Data of Each Condition

Table 5.2 Names and descriptions for each parameter used in the version of PEPPR implemented here.

Purpose	Parameter	Description
Encoding	ϕ_s	Scaling of primacy gradient in learning new context-feature associations
	ϕ_d	Rate of decay of primacy gradient
	γ_{fc}	Strength of new feature-context associations
	γ_{cf}	Strength of new context-feature associations
	β_{enc}	Rate of context drift during encoding
Retrieval	β_{rec}	Rate of context drift during recall
Decision Process	θ_s	Scaling of probability of recall failure
	θ_r	Rate at which the likelihood of recall success decreases with additional output position
	τ	Sensitivity to differences in activation at retrieval for luce choice rule
Cognitive Control	$\beta_{PEPPR_{encoding}}$	Rate of context drift when goals are encoded
	$\beta_{PEPPR_{retrieval}}$	Rate of context drift when goals are retrieved
	ϵ	Strength of associations between non-item features (location and dish) and goals

dishes. The model was fit to data from the free recall condition using a genetic algorithm that ran for 2,500 generations with a total of $k = 12$ free parameters. Model fit was measured using root mean square deviation (RMSD). For additional details on the methods for model fitting and the resulting parameter set, see Appendix A.

Next, using the parameter values obtained from fitting to the free recall condition, data were simulated for two sets of 60,000 simulated subjects, each recalling 2 lists. For the first set of simulations, the location-focus goal node was activated before encoding and before retrieval for each list. In the second set of simulations, the dish-focus goal node was activated before encoding and before retrieval for each list.

Modeling Results

Simulated data from the model fit to the free recall condition is presented alongside the behavioral data in Figure 5.12. The model fit overall recall as well as the general shape of the TCE, with a much higher bias for near-lag transitions, particularly in the forward direction. The model

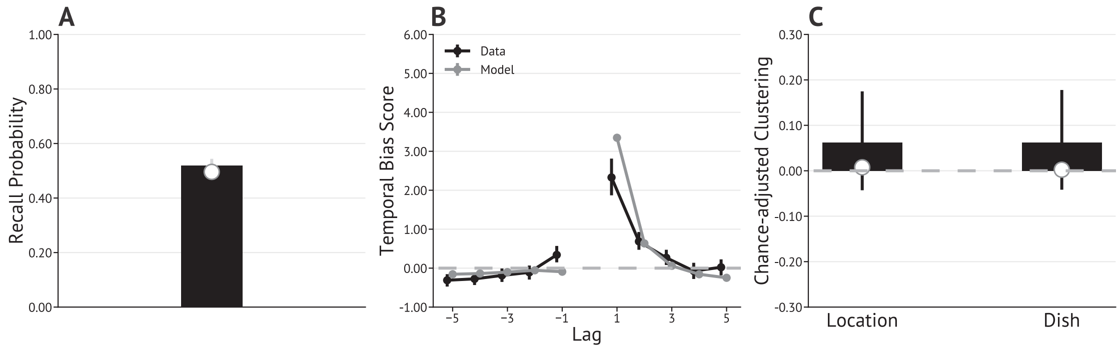

Figure 5.12 Behavioral data for the free recall condition plotted alongside simulated A) recall probabilities, B) temporal bias scores, and C) chance-adjusted location and dish category clustering scores from the Post-Encoding Pre-Production Reinstatement Model (PEPPR) modified to include a representation of task goals. Error bars represent bootstrapped 95% confidence intervals.

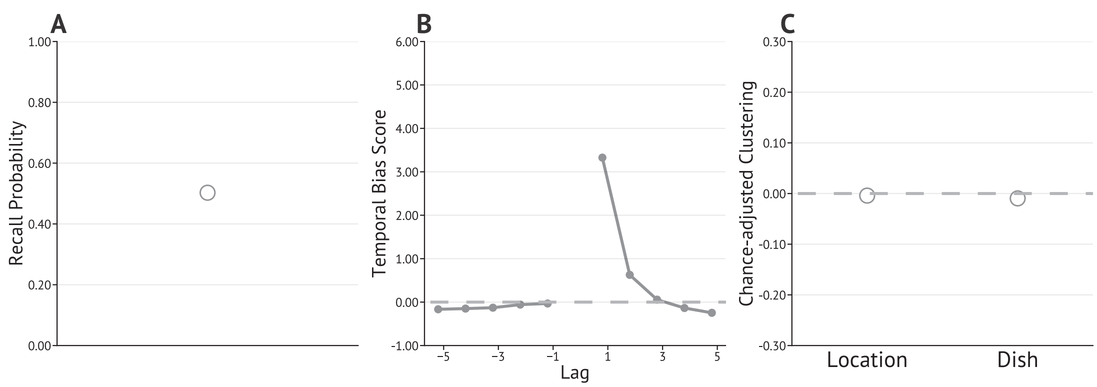

Figure 5.13 Simulated A) recall probability, B) temporal bias scores, and C) chance-adjusted location and dish category clustering scores for the location-focus condition from the Post-Encoding Pre-Production Reinstatement Model (PEPPR) modified to include a representation of task goals.

also accurately captured the null clustering effects for organization by location or dish category. The parameter set from this fit was used to simulate predictions for the location and dish clustering conditions. For these simulations, the only difference was that the location-focus or dish-focus goal node was activated during encoding and retrieval.

Simulated data for the location-focus condition is presented in Figure 5.13 and for the dish-focus condition in Figure 5.14. In both cases, the model essentially predicts no effect of task goals. The model predicts similar levels of recall, temporal contiguity, and category clustering for all three

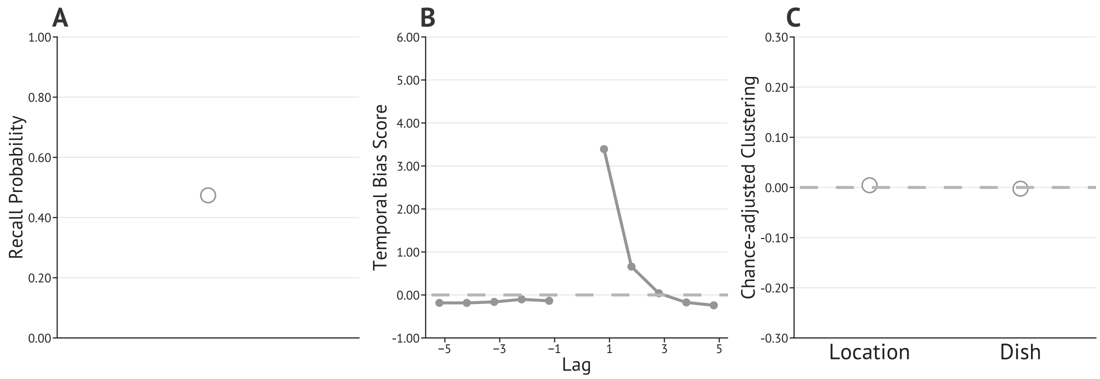

Figure 5.14 Simulated A) recall probability, B) temporal bias scores, and C) chance-adjusted location and dish category clustering scores for the dish-focus condition from the Post-Encoding Pre-Production Reinstatement Model (PEPPR) modified to include a representation of task goals.

conditions. Importantly, both simulations display a null category clustering effect for both location and dish. Although this is consistent with the behavioral data, the goal of these simulations was to test if representing goals and relevant features in context was sufficient to produce a category clustering effect for the task-relevant categories. In this respect, the simulation was not successful because the activation of the goal node did not result in any changes in clustering for the simulated data.

Interim Discussion

Participants did not cluster their recalls by either the associated location or the associated dish in any condition. Although such a lack of clustering might be expected in the free recall condition, it is somewhat surprising that even when participants were instructed to remember the items *for a shopping trip to the three locations* or *for cooking the three dishes*, they did not organize memory search based on these associations. There are multiple possible explanations for these null category clustering effects.

In Chapter 4, I found that participants' strategies during encoding affected some aspects of recall organization, but goals during retrieval were much more influential. However, the kinds of associations being considered in Chapter 4 were temporal associations, which appear to be largely automatically encoded, and semantic associations, which rely on pre-existing knowledge and do not rely on the formation of new associations during encoding. In contrast, it is not clear if the location and dish associations are encoded automatically, and they do not rely on pre-existing knowledge. The availability of location and dish associations during retrieval may depend on intention during encoding. Alternatively, even if the location and dish associations were encoded, participants may have opted not to use them to guide retrieval. The question of if the location and dish associations were encoded was addressed by conducting a second experiment that directly tested participants' memory for the location and dish associations.

Experiment 4b

Methods

The methods for Experiment 4b were identical to those described above for Experiment 4a with two exceptions: the number of conditions and the test format. Because the aim of this experiment was to test if location and dish associations were encoded when they were task-relevant (in the location-focus and dish-focus conditions respectively), there was no free recall condition. Participants received the same instructions as in Experiment 4a, leading them to expect a free recall test. However, after each list and a reminder of their goal, participants' memory for the location or dish associated with each item was instead directly tested with a multiple-choice recognition test.

For the multiple-choice test, one of two prompts appeared at the top of the screen, depending on condition. Participants in the location-focus condition were asked "Which store was paired with this ingredient when you studied the list?" Participants in the dish-focus condition were asked "Which dish was paired with this ingredient when you studied the list?" Importantly, participants in each condition were tested on their memory for only the associations that were relevant to their assigned goal (location or dish). An item from the previously studied list appeared onscreen with three answer options below it. For the location-focus condition, the answer options were SUPERMARKET, FARMER'S MARKET, and SPECIALTY STORE. For the dish-focus condition, the answer options were APPETIZER, MAIN DISH, and SIDE DISH. Each answer option was numbered (1, 2, and 3), and participants indicated their answer by pressing the number key corresponding to their choice. The items were tested in random order, and the order of the answer choices was randomized for each item.

Participants

Given that the goal of Experiment 4b was to detect if recognition of location and dish associations was above chance, a smaller sample was collected than in Experiment 4a. I set out to collect data from 122 participants per condition, which would provide 95% $1 - \beta$ power to detect recognition above chance with an effect size of $d = 0.3$. In total, 249 Michigan State undergraduates who had not participated in Experiment 4a completed the experiment on their personal computers. Twelve participants were excluded for reporting on the post-task questionnaire that there was a reason their data should be excluded (e.g., being distracted during the task or not understanding the instructions). Five additional participants were excluded for attempting the experimental task multiple times. In the final sample of 232 participants, 188 identified their gender as female, and the average age was 19.4 years (SD = 1.8).

Results

Recognition Performance

Recognition performance is displayed in Figure 5.15. Recognition performance was measured for each participant as the proportion of correct answers on the multiple-choice test and can be compared to the chance performance level of 0.33. Chance performance represents the proportion of correct answers that would be expected if participants were randomly guessing one of the three answer choices. Performance above chance would indicate that participants did encode the location and dish associations, while performance at chance would indicate that the associations were not initially encoded. Each condition was only tested on the associations that were *task-relevant*. The location-focus condition was tested on their memory for each item's associated location, and the dish-focus condition was tested on their memory for each item's associated dish.

Recognition performance was above chance in the location-focus condition (M = 0.46, SD = 0.50), as indicated by the 95% confidence intervals not crossing chance. However, recognition performance was no different from chance in the dish-focus condition (M = 0.34, SD = 0.47). These results provide some insight into the question of if the location and dish associations are encoded but not used to guide retrieval or if they are not encoded at all. Participants did encode location associations, at least to some extent. However, they were not able to retrieve the dish type associated with each item as would be expected if these associations were never encoded or were encoded in a way that made later retrieval difficult.

Performance on the recognition test was similar for both List 1 and List 2, even though participants were surprised by the multiple-choice test at the end of List 1 and (presumably) expected the multiple-choice test on List 2. Mean recognition accuracy for each condition is listed separately for List 1 and List 2 in Table 5.3.

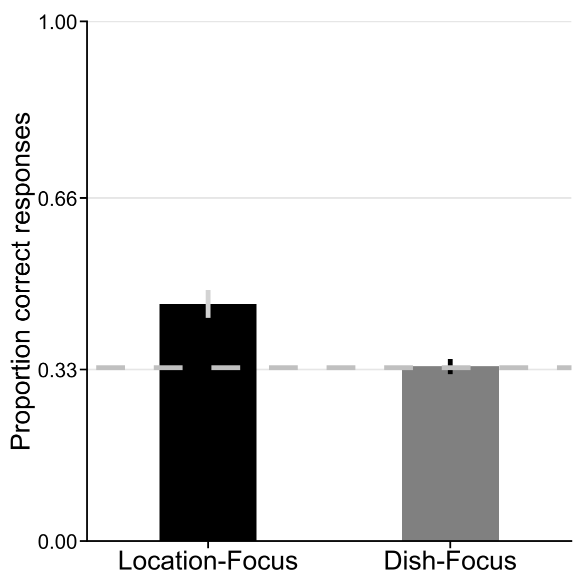

Figure 5.15 Proportion of correct responses to the multiple-choice recognition test. Participants in the location-focus condition were tested on their memory for the location associated with each item, and participants in the dish-focus condition were tested on their memory for the dish associated with each item. Chance performance (0.33) is indicated with a dashed line. Error bars represent bootstrapped 95% confidence intervals.

Self-Reported Strategy Use

In a post-task questionnaire, participants were asked to report any strategies they used during study. Participants were not asked to report their use of recall strategies, since there was no recall test in Experiment 4b. The proportion of participants indicating that they used each encoding strategy is presented in Figure 5.16. As in Experiment 4a, most participants reported using shallow

Proportion Correct Responses on Recognition Test
Table 5.3 Chance (random guessing) would result in 0.33 of responses being correct.

Condition	Mean Proportion Correct Responses (SD)	
	List 1	List 2
Location-Focus	0.45(0.50)	0.46(0.50)
Dish-Focus	0.33(0.47)	0.34(0.47)

strategies, such as "Reading each word as it appeared" (64.7%) and "Repeating the words as much as possible" (57.1%).

In the location-focus condition, 40.3% of participants reported "Focusing on the location where each item could be purchased," while only 8.8% of those in the dish-focus condition reported strategically encoding location associations. Fewer than half of participants focused on location, even though it was task-relevant; this is consistent with the recognition performance, which was above chance but still relatively low. In comparison to Experiment 4a, where only 8.9% of participants reported focusing on location associations during encoding, a much larger proportion of participants used location-based encoding strategies.

A smaller proportion of participants reported "Focusing on the dishes each item would be used to cook." Only 23.0% of participants in the dish-focus condition and 10.9% of participants in the location-focus condition reported prioritizing the dish associations. As with the location-focus condition, the proportion of dish-focus participants reported focusing on the dish associated with each item during study was higher in Experiment 4b than in Experiment 4a (where only 8.6% of participants reporting prioritizing dish associations during study). This difference could be due to participants in Experiment 4b changing their strategy for the second list. Experiments 4a and 4b were identical until the test after the first list, where participants in Experiment 4a completed a free recall test and participants in Experiment 4b completed a multiple-choice recognition test. Presumably, participants in Experiment 4b noticed that they were being directly tested on dish (or location) associations and thus may have focused more on the dish (or location) associated with each item while studying List 2, even though this had no effect on overall performance. In any case, the results of the self-reported strategy use in Experiment 4b are consistent with the low levels of recognition performance, especially the lower performance in the dish-focus condition.

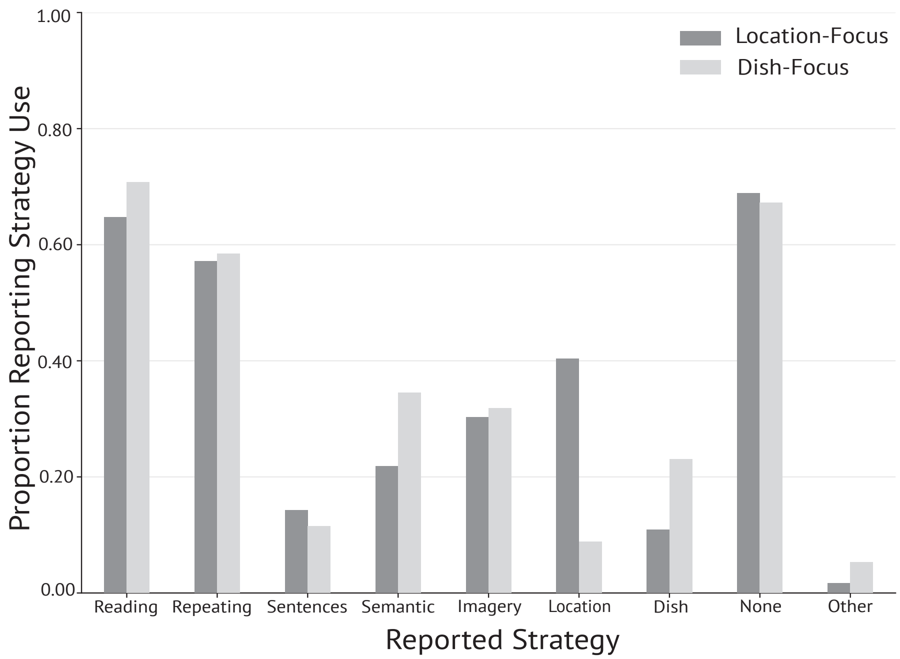

Figure 5.16 Encoding strategies reported by participants in Experiment 4b. The proportion of participants reporting using each strategy is plotted for each condition. Since participants could select multiple answers, the proportions for each condition may sum to > 1.0.

Interim Discussion

Experiments 4a and 4b examined patterns of recall organization in a more complex situation than previous experiments: when to-be-remembered items were related along multiple associative dimensions and participants were provided an external goal that encouraged organization along one of those dimensions. In Experiment 4a, participants displayed a robust TCE regardless of their task goals, consistent with retrieved context models. However, organization along other associative dimensions was minimal even when those associations were task-relevant. When participants were tested directly on their memory for the location or dish associated with each item in Experiment 4b, recognition of the task-relevant association was low in the location condition ($M = 0.46$) and at chance for the dish condition ($M = 0.34$). Regardless of task goals, participants primarily organized their recalls based on temporal order, not based on other task-relevant associations.

These findings further support some of the core claims of retrieved context models: temporal order information is an integral part of episodic memories, and temporal context is encoded whenever new memories are formed. If temporal associations drive recall organization in free recall of random word lists merely because they are the only useful associations available (e.g., Hintzman, 2011), then the TCE should be greatly reduced or even abolished when other concrete associations are available and those associations do not parallel the order of items during encoding. This was not the case. A strong TCE was observed even when participants could organize their recalls based on the location in which they could purchase the items or the dish the items would be used to make. Furthermore, different goals did not affect the degree of temporal organization. The TCE also predicted recall success in all three conditions, suggesting memory for items and memory for their order were tightly bound together. These results are all consistent with the predictions for temporal contiguity made by retrieved context models.

Surprisingly, although they showed clear evidence of temporal organization, participants did not organize memory search based on each item's associated location or dish even when those associations were relevant to their assigned goals. One possible reason for the absence of category clustering is that participants were not attending to the task, which is of particular concern when participants complete the experiment using their personal computers (see the pilot studies described in Chapter 4). However, it is unlikely that poor attention to the task is responsible for the lack of clustering observed here. Participants who reported not attending to the task were excluded, and participants who were included did well on the task; average recall performance was high relative to Experiments 1-3 and other online experiments with the same sample (e.g., Healey & Uitvlugt, 2019). And even the highest-performing participants did not display significant category clustering (see Figures 5.6, 5.7, and 5.8).

Another possibility is that participants did attend to the task in general but not to the location and dish associated with each item. This explanation assumes that encoding the location and dish associations was not automatic but required intention and effort. If participants did not intentionally encode the location and dish associated with each item, then those associations would not be available to guide later recall. The results of Experiment 4b support this explanation. When participants in the dish-focus condition were directly tested on the dish associated with each item, performance was no better than chance. Memory for the location associated with each condition was above chance for the location-focus condition, but performance was still fairly low (participants identified the correct location for 46% of items on average; chance performance was 33%). If participants were unable to correctly identify the location or dish associated with each item on a recognition test, it is certainly possible that these associations were never encoded, or encoded only weakly.

Participants may have not encoded the location and dish associated with each item due to the limited amount of time they had during encoding. Based on previous experiments conducted in our lab, the 5 s encoding period was sufficient to study three words. However, this does not necessarily mean that participants devoted equal time and attention to all of the words on the screen. Assuming attention is a limited resource (e.g., Kahneman, 1973), focusing attention only on task-relevant features should improve later memory performance (Benjamin, 2007). Participants could have allocated their limited time and attention to learning the ingredients, since the primary goal in all conditions was to remember the ingredients, rather than to also learning the location and dish associated with each item. It is also important to consider the influence that the specific materials used in this experiment may have had on the results. More concrete and discrete categories (e.g., specific stores the participants would recognize) might be more salient and therefore better remembered and utilized by participants than the general categories used in the present experiment. Such a possibility should be explored in future work.

With regards to automatic and controlled influences on memory, encoding and retrieval of ad-hoc category associations may be based primarily on strategic control processes. In Experiment 4a, few participants intentionally focused on location or dish information during either encoding or retrieval, and in the absence of these control processes, these associations had little influence on recall organization. In contrast, the temporal organization observed in Experiment 4a could be a result of a combination of automatic and controlled processes, especially if participants were influenced by their prior experience with free recall tests. In memories outside the lab, certain kinds of associations are typically encoded, such as spatial location (Gibson et al., 2021; Kowialiewski et al., 2022; Miller et al., 2013; Pacheco & Verschure, 2018), while others are not, like the color of studied items (e.g. Hong, Polyn, & Fazio, 2019). However, it is unclear the degree to which task goals determine the encoding of various kinds of associations (e.g., if spatial location is included as a variable in a task, it is likely to be task-relevant).

Developing a computational model that can account for the influence of task goals on recall organization is an important next step in producing a model of memory that can account for both the automatic and controlled influences on organization that have been found previous experiments. Even though there was no effect of task goals on recall organization in this experiment, I was able to test one method of implementing task goals in retrieved context models: representing task goals in context in the same way that to-be-remembered items and other features were represented. Although this new model implementation successfully fit to recall success and organization in the free recall condition, activating the location-focus or dish-focus goal failed to influence organization in model simulations. More work is needed to test alternate ways of representing control processes in these models. Potential future directions for model development are considered in the General Discussion.

These results demonstrate that temporal associations influence memory search even when other associations are available and task-relevant, as is often the case for episodic memories formed outside the lab. There was a strong TCE in all conditions, even when participants' assigned goal encouraged them to organize their recalls based on each item's associated location or dish, which should have reduced temporal contiguity. In contrast, participants did not show evidence of clustering by location or dish in any condition, and participants' self-reports of strategy use indicate that few participants strategically prioritized location or dish associations during encoding or retrieval. In the absence of strategic control processes to encourage encoding and recall of the dish and location information, participants did not remember the dish or location associated with each item well, suggesting that these associations are primarily learned through strategic control processes. Of all the novel associations available to participants, temporal context was the most influential on recall organization and predicted recall success, consistent with the retrieved context model account.

CHAPTER 6

GENERAL DISCUSSION

Episodic memories are often recalled in the order they were originally experienced, and Experiments 1-4b demonstrate that this temporal organization is influenced by both automatic and controlled mechanisms. These experiments tested the predictions of retrieved context models, which emphasize the role of automatic context-based mechanisms, and accounts that emphasize the role of control processes in generating the temporal contiguity effect (TCE). This effect is a useful tool for distinguishing the effects of automatic and controlled processes because there is evidence that temporal information is automatically encoded, yet the size of the TCE is modulated by strategic control processes.

Four experiments were designed to answer empirical questions regarding the interaction of automatic and controlled processes on recall organization: 1) Is temporal information automatically encoded and automatically retrieved? 2) How does assigning different tasks during encoding affect recall organization? 3) To what extent do strategic control processes during encoding determine the availability of information for later recall organization? and 4) Is the TCE present even in the presence of external, non-temporal goals? In addition, I tested the models' ability to account for other forms of recall organization and the influence of task goals by fitting a computational model to the data from Experiments 3 and 4a.

In the following sections, I provide a summary of the theories motivating this work and the experiments which examine each question. Then, I discuss the evidence for automatic and controlled influences on the TCE and the implications for memory theories. Finally, I consider open questions raised by some of these findings and future directions for developing computational models that combine the automatic mechanisms of retrieved context models with mechanisms from neural models of cognitive control.

Summary

Retrieved context models (e.g., Howard & Kahana, 1999; Howard et al., 2015; Lohnas et al., 2015; Polyn et al., 2009a) provide a well-specified account of the automatic context-based mechanisms underlying episodic memory. Under these models, new episodic memories are created when an association is formed between a representation of the item being studied and the current state of temporal context. As each item is processed, it brings to mind its pre-existing associations, updating context to reflect these newly activated associations. When the next item is studied, it automatically becomes associated with this updated context and then brings to mind its own existing associations, further updating the context. When the next item is studied, the previous context representation is not completely erased. Rather, context is a blend of the previous context and the new context. As a result, items experienced relatively closer together in time are associated with more similar states of context. When an item is retrieved, it reinstates its associated context, providing a good cue for other items experienced nearby in time because they were associated with similar states of context during study. Thus, these models naturally predict a TCE in almost any circumstance.

An alternate account of the TCE attributes the effect primarily to strategic control processes, intentional strategies used by participants during encoding or retrieval that encourage recalling items in order. A strict control processes account is unable to account for the TCE found when the use of temporal strategies is unlikely, such as following incidental encoding (e.g., Healey, 2018; Mundorf et al., 2021). Nonetheless, there is evidence that the TCE is influenced by strategic control processes; the size of the effect varies across individuals and situations, particularly when the capacity for cognitive control is limited (e.g., Healey & Kahana, 2016; Healey & Uitvlugt, 2019; Long & Kahana, 2017; Spillers & Unsworth, 2011). Although retrieved context models provide a well-specified account of the automatic mechanisms underlying the TCE, they lack explicit mechanisms for control processes. Neural models of cognitive control provide guidance on how strategic control processes might be implemented in these models: representing task goals as a part of context and allowing task goals to bias context representations at encoding and retrieval. The purpose of Experiments 1-4b was to test the predictions of both accounts and disentangle the effects of automatic and controlled processes on recall organization.

Experiment 1 directly tested the retrieved context models' prediction that when an item is retrieved, its associated temporal context is automatically retrieved as well. In this experiment, participants completed an implicit memory test where they read words aloud, and implicit memory was gauged by faster responses to words on their second presentation. Simply reading a word was sufficient to retrieve its temporal context and prime responses to other items studied nearby in time. These results demonstrate temporal information is not only automatically encoded (as evidenced by the TCE observed in the surprise free recall task; see also Healey, 2018; Mundorf et al., 2021) but also demonstrate that it is automatically retrieved to some extent.

The subsequent experiments focused on the influence of strategic control processes at encoding and retrieval on the TCE. Experiment 2 tested the prediction of retrieved context models that deep processing (processing items for their meaning) should increase item-specific processing and thereby increase the TCE against other theoretical predictions. As predicted, deep processing did increase temporal organization relative to shallow processing. However, both recall and the TCE peaked in the no-task condition. Assigning any task during encoding interfered with temporal organization, suggesting that control processes also influence the TCE.

Experiment 3 investigated the degree to which control processes during encoding determine the kinds of associations that are later available to guide recall. If control processes predominately operate at encoding by directing attention to task-relevant features (Benjamin, 2007; O'Reilly et al., 1999; Summerfield, 2006; Wagner, 2002), then the amount of temporal information available to guide recall should depend on participants' encoding strategies. If, instead, temporal information is always automatically encoded, differences in recall organization should depend primarily on differences in retrieval strategies. In Experiment 3, instructions to ignore temporal information did not abolish the TCE, consistent with retrieved context models, but there was also a clear effect of strategic control processes. Both the TCE and semantic organization were modulated primarily by participants' assigned retrieval strategies. This demonstrates that, at least with regards to temporal and semantic contiguity, cognitive control influences recall organization at retrieval.

Experiments 4a and 4b examined retrieved context models' prediction regarding the TCE, particularly that the effect should occur in more intricate scenarios involving the presence of other types of associations that were both available and task-relevant. In these experiments, participants were assigned different goals ("Try to remember the words for shopping at different stores", "Try to remember the words for when you cook the dishes") that encouraged them to focus on either the location or dish presented with each item at encoding. As predicted by retrieved context models, a TCE was observed in all conditions and was unaffected by participants' assigned goals. Surprisingly, however, participants did not organize their recalls based on location or dish even when those associations were task-relevant. A direct test of memory for the location or dish associated with each item revealed poor encoding of most of the item-location and item-dish associations. In these experiments, temporal organization was preferred even when the other associations were task-relevant.

Evidence for Both Automatic and Controlled Effects on the TCE

To distinguish the effects of automatic and controlled processes on the TCE, the present results can be considered in light of Hasher and Zacks's (1979) framework. Based on this framework, if the TCE is due to automatic mechanisms, temporal organization should occur even in the absence of intentional study or retrieval, not interfere with other processes, and be consistent across individuals and situations. If the TCE is instead a result of controlled processes, then the effect should be eliminated when participants are not intentionally trying to encode or retrieve temporal information, be reduced by other interfering processes, and occur only for some individuals and in some situations.

Does the TCE Occur in the Absence of Intention to Encode or Retrieve?

One of the clearest indications that the TCE is a result of automatic mechanisms is that the effect occurs even when encoding and retrieval are unintentional. In my own previous work, I found a small but significant TCE under incidental encoding conditions (Mundorf et al., 2021, see also; Diamond & Levine, 2020; Healey, 2018; Moreton & Ward, 2010; Pathman et al., 2023; Uitvlugt & Healey, 2019). This finding was replicated in Experiment 1, where a strong TCE was observed in a surprise free recall test. In addition, instructions to completely ignore temporal order during encoding had no effect on the magnitude of the TCE in Experiment 3. These findings support the claim that temporal information is automatically encoded and are consistent with retrieved context models.

Experiment 1 provides direct evidence that temporal information is also automatically retrieved. In this experiment, participants displayed significant associative repetition priming on an implicit memory test. That is, when they read a word, its temporal context was automatically retrieved and primed responses to other items studied nearby in time. Because in this task retrieval was truly unintentional, this experiment provides a more exact test of automatic retrieval than previous work where temporal information was merely unhelpful (Davis et al., 2008; Osth & Fox, 2019). Similarly, in Experiment 3 temporal information influenced recall even when participants were instructed to ignore it completely, as would be expected if temporal information was automatically retrieved. However, the TCE was significantly greater for those participants instructed to focus on temporal order during retrieval compared to those who were not. Together, the results of Experiments 1 and 3 support the claim that temporal information is encoded and retrieved automatically while also pointing to the critical role of strategic control processes in determining the extent to which temporal information guides memory search.

Does the TCE Interfere with Other Processes?

The second of Hasher and Zacks's (1979) criteria for distinguishing between automatic and controlled mechanisms has to do with interference: does the TCE compete with other tasks for cognitive resources? Previous work addressing this question resulted in mixed findings (Long & Kahana, 2017; Murphy & Castel, 2021), likely because of differences in the kinds of additional tasks participants had to complete during encoding. This question is addressed by Experiment 2. A TCE was present regardless of whether participants were assigned to complete a deep processing task, a shallow processing task, or no task during encoding, as predicted by retrieved context models. However, both recall and temporal contiguity were greatest when participants were not assigned an encoding task. That is, assigning an additional task during encoding did reduce the TCE, as would be expected if strategic control processes contributed to the TCE (Long & Kahana, 2017). Here again, the results are consistent with both the automatic mechanisms of retrieved context models and the influence of strategic control processes.

Is the TCE Consistent Across Individuals and Situations?

Memory phenomena resulting from automatic mechanisms should also be consistent across individuals and situations. A TCE was consistently observed in the average data across all experiments, and some patterns of individual differences were remarkably consistent. In all three experiments where individual differences were examined, participants with greater temporal contiguity also had higher levels of recall, regardless of encoding or retrieval instructions (see also Polyn et al., 2011; Sederberg et al., 2010; Spillers & Unsworth, 2011). These patterns are consistent with retrieved context models' prediction that the TCE and recall should be tightly linked, since in these models temporal context is an integral part of all episodic memories.

Yet, the presence of individual differences at all suggests control processes may be at work. As is evident in the scatter plots of temporal, semantic, and category-based organization for Experiment 4, some participants displayed high levels of organization, while others did not. In addition, the kinds of information that were associated with better recall varied depending on participants' intentional strategies. In Experiment 3, the correlation between the TCE and recall was strongest when participants were assigned a temporal focus during both encoding and retrieval. Semantic contiguity was also positively correlated with recall, but only in conditions assigned a semantic test strategy.

The TCE was not eliminated in any of the experiments, including when tasks were assigned during encoding (Experiment 2), other useful associations were available (Experiments 3 and 4a), or participants were explicitly instructed to ignore temporal order information during encoding and retrieval (Experiment 3). These findings are consistent with retrieved context models and suggest that the TCE is, at least in part, a result of automatic mechanisms. Yet, the size and shape of the effect varied, as would be expected if temporal organization is also influenced by strategic control processes. The TCE was greatest under conditions that allowed for the greatest engagement of temporally-based strategies, such as the no-task condition in Experiment 2 and the conditions assigned temporal retrieval strategies in Experiment 3.

A detailed look at the recall dynamics reveals another pattern present across experiments. Focusing on temporal order during encoding resulted in an especially high bias for making +1 lags. This forward asymmetry is a prominent characteristic of the TCE observed in a typical free recall test. Yet, forward asymmetry is greatly reduced when participants are less likely or less able to engage in order-based strategies during encoding. Previous work has found individual differences in cognitive control ability are related to differences in the TCE, particularly the bias for *lag* = +1 transitions. For example, older adults, who typically have a reduced capacity for cognitive control (Hasher et al., 2007; West, 1996), display a smaller and more symmetric TCE than younger adults, with the main difference being a reduced probability of making *lag* = +1 transitions (Diamond & Levine, 2020; Healey & Kahana, 2016; Kahana et al., 2002). Direct manipulations that decrease the likelihood that participants will engage in order-based encoding strategies also reduce forward asymmetry (Healey & Uitvlugt, 2019; Mundorf et al., 2021). From the perspective of a strategic control processes account, this is unsurprising. If forward asymmetry is a result of intentional encoding strategies like linking the words together to form a story, a strategy commonly reported in free recall tasks (Bouffard et al., 2018), then when participants are unable to engage in those strategies effectively, the forward bias should be reduced or even disappear.

This pattern is also evident in the experiments presented here. In Experiment 1, when participants were given a surprise free recall test and thus had no reason to engage in order-based strategies during encoding, the TCE was symmetrical with no greater bias for recalling items in forward order (see also Mundorf et al., 2021). The TCE was also symmetrical for the deep processing condition in Experiment 2. The encoding task for the deep condition was most time-consuming and could have interfered with TCE-generating encoding strategies to a greater extent than the shallow processing condition, where there was a slightly higher bias for +1 lags, or the no-task condition, where forward asymmetry was most pronounced. In Experiment 3, encoding strategy did influence the bias for making *lag* +1 transitions even though it did not affect overall temporal contiguity. A focus on temporal order at both encoding and retrieval resulted in the greatest forward asymmetry. Across experiments, participants who were more able to engage in order-based control processes during encoding displayed a much higher bias for +1 lags in particular.

Within retrieved context models, forward asymmetry can be explained using existing automatic mechanisms. In the model, when an item is studied, it automatically forms an association with the current state of context (experimental associations) and then reinstates its own pre-experimental associations, which are then integrated into context (for a visual example, see Figure 1.3). The relative influence of the pre-experimental and experimental contexts is controlled by a model parameter. If pre-experimental associations are more influential, then the model predicts a greater bias for making forward transitions. This is because when an item is recalled and reinstates its associated contexts, the experimental associations are an equally good cue for both items studied immediately before and immediately after the just-studied item. In contrast, the pre-experimental associations provide a good cue only for items studied after the just-studied item. This is because the pre-experimental context is only associated with items that follow the just-presented item, not those presented earlier in the list. Yet, a representation of control processes that would allow retrieved context models to predict *a priori* that forward asymmetry should be greater in some conditions is absent from these models. As with the broader patterns of recall organization, forward asymmetry in the TCE would be best explained by a combination of the automatic mechanisms of retrieved context models and a representation of the effects of control processes, like chaining items together during encoding.

Open Questions

Retrieved context models provide a comprehensive explanation of the relationship between memory success and organization. Yet, two challenges limit the generalizability of these models. First, the TCE has been studied primarily in free recall of unrelated words with the goal of recalling as many words as possible. This task is ideal for isolating the effect of temporal associations. However, there is still some question of whether these models can explain organization outside the lab where events are related along multiple associative dimensions and goals may be more complex. Another limitation is that retrieved context models rely only on automatic mechanisms despite evidence that both automatic and controlled mechanisms interact to produce patterns of recall organization. I plan to work towards addressing both of these limitations in future work, using the experiments and computation modeling conducted here as a baseline.

Recall Organization Along Multiple Dimensions

The goal of Experiment 4a was to test if temporal information guides memory search in a situation more complex than a typical free recall task where temporal associations are the only useful associations that are available. Not only was the TCE not eliminated, but participants actually preferred temporal organization to organization based on the location or dish presented with each item. This further supports the claim that the TCE is not merely a product of an artificial laboratory task. However, the absence of any kind of clustering by location or dish, even when those associations were relevant, leaves unanswered other questions about how task goals affect the balance between different kinds of associations.

For example, does the effect of task goals depend on the *kinds* of associations available? Experiments 3 and 4 are both concerned with the effects of goals on temporal contiguity when other kinds of associations are available. Yet, the results of these two experiments appear to lead to different conclusions. In Experiment 3, task goals at retrieval had a large influence on both temporal and semantic organization; in Experiment 4, task goals had no effect on organization of any kind. An important difference between these two experiments is the kinds of non-temporal associations that were task-relevant. The semantic associations in Experiment 3 were based on *pre-existing* knowledge. In contrast, organizing by location or dish in Experiment 4 would have required participants to encode *novel* associations. Forming new, non-temporal associations may have required more effort than participants were willing or able to expend. It is important for future work to directly compare how recall organization differs for novel compared to pre-existing associations to better understand the underlying cause of the different effects of task goals in Experiments 3 and 4a.

Another question raised by the results of Experiments 4a and 4b is how far the results extend to other materials and situations. Would a similarly strong TCE be observed with other associative dimensions that are more salient or easier to conceptualize? Some features, like text color, do not influence recall organization (Hong, Polyn, & Fazio, 2019). Previous work has found evidence of clustering by other features like spatial location (Clark & Bruno, 2021; Gibson et al., 2019; Kowialiewski et al., 2022; Miller et al., 2013; Pacheco & Verschure, 2018; Robin et al., 2016), reward (Murphy & Castel, 2021; Stefanidi & Brewer, 2015), and emotion (Long et al., 2015; Siddiqui & Unsworth, 2011) in tasks where these associations are task-relevant. However, the relevance of these associations has not been manipulated in a single experiment. Such an experiment would be informative with regards to the relationships between organization along different associative dimensions.

Representing Strategic Control Processes in Retrieved Context Models

Future experimental work where task goals determine recall organization will pave the way for further development of how mechanisms for cognitive control can be integrated into retrieved context models. Neural models of cognitive control emphasize the roles of the PFC, the hippocampus, and communication between these two brain regions in cognitive control of memory (Eichenbaum, 2017a; Polyn & Kahana, 2008). Such models provide guidance on how strategic control processes might be implemented in retrieved context models: by representing task goals as a part of context and allowing task goals to bias context representations at encoding and retrieval.

The model used to simulate the effect of goals on recall organization in Chapter 5 represents only one possible way control processes might be represented in retrieved context models. To-be-remembered items, task goals, and task-relevant features were all represented in context in the same way. Task goals influenced retrieval by making task-relevant features more likely to be retrieved and then used as a cue for recalling items. However, there are alternative ways in which task goals could be represented. For example, Polyn et al. (2009a) represented task goals as a part of context but distinguished between *temporal* context, which included a record of recently presented items, and *source* context which represented task goals. This kind of representation could allow for task goals to be maintained in the face of distractors and other events (Polyn & Kahana, 2008), a core feature of cognitive control (Duncan, 2010; Hazy et al., 2006; Miller & Cohen, 2001; O'Reilly et al., 1999; Wagner, 2002) that is missing from the model proposed in Chapter 5. However, it is unclear if this kind of representation could explain how task goals bias processing.

Mechanisms through which strategic control processes can bias processing at both encoding and retrieval are also important areas for model development. Neural evidence points to communication between the PFC and hippocampus as a critical component in explaining cognitive control of memory (Eichenbaum, 2017a; McClelland et al., 1995; Moscovitch, 1992). This communication is represented in PEPPR with goals during the retrieval process, where activating the goal representation led to an increased likelihood of recalling the task-relevant associations, based on pre-existing associations between the features and the goal representation. The PEPPR with goals model does not, however, include mechanisms for biasing processing during encoding. Task goals do influence the kind of information that is encoded (O'Reilly et al., 1999; Summerfield, 2006; Wagner, 2002). In Experiment 3 participants' goals during encoding did not affect temporal or semantic organization. However, task goals may be more important for encoding novel, non-temporal associations, such as the location and dish associations in Experiment 4a.

Another aspect of cognitive control that is underdeveloped in current implementations of retrieved context models is how strategies are learned and developed. Behaviorally, participants are often able to pick up on the associations in a list and to engage in strategic control processes based on what kinds of associations are available even when they are not instructed to adopt a particular strategy (e.g., semantic strategies in lists with a semantic structure; see Hong et al., 2022; Polyn et al., 2011). A comprehensive model of memory should also be able to model this process. It remains to be seen how the learning of goals could be implemented in retrieved context models. Developing this kind of learning process may require the implementation of a connectionist learning network, such as in the model of semantic organization developed by Becker and Lim (2003). Such a step is crucial to develop a model that can represent every part of the learning and retrieval process.

Conclusion

Across experiments, the TCE was found to be a result of both automatic mechanisms and strategic control processes. The TCE was observed under a myriad of situations, as predicted by retrieved context models, including in an implicit memory test and in free recall when participants were instructed to ignore temporal order information and when other associations were available. However, the size and shape of the TCE were also affected by participants' ability to engage in order-based strategies. A full explanation of these results requires an integration of the automatic mechanisms of retrieved context models and control processes from neural models of cognitive control. Control processes' influence on recall organization can be represented in retrieved context models through representations in context and by weighting associations. However, current versions of the models may not be able to account for high levels of organization along non-temporal dimensions. Future work should continue to explore ways in which the principles of neural models of cognitive control can be integrated with retrieved context models to provide a comprehensive, detailed theory of episodic memory.

BIBLIOGRAPHY

Alexander, M., Stuss, D., & Fansabedian, N. (2003). California Verbal Learning Test: Performance by patients with focal frontal and non-frontal lesions. *Brain*, *126*(6), 1493–1503. https://doi.org/10.1093/brain/awg128

Alkire, M., Haier, R., Fallon, J., & Cahill, L. (1998). Hippocampal, but not amygdala, activity at encoding correlates with long-term, free recall of nonemotional information. *Proceedings of the National Academy of Sciences, USA*, *95*(24), 14506–14510. https://doi.org/10.1073/pnas.95.24.14506

Audacity Team. (2018). Audacity®: Free audio editor and recorder [Audacity® software is copyright © 1999-2022 Audacity Team. The name Audacity® is a registered trademark.].

Averell, L., Prince, M., & Heathcote, A. (2016). Fundamental causes of systematic and random variability in recognition memory. *Journal of Memory and Language*, *88*, 51–69. https://doi.org/10.1016/j.jml.2015.12.010

Ayçiçegi-Dinn, A., & Caldwell-Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism*, *12*(3), 291. https://doi.org/10.1017/s1366728909990125

Baddeley, A. D. (1978). The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review*, *85*(3), 139–152. https://doi.org/10.1037/0033-295x.85.3.139

Becker, S., & Lim, J. (2003). A computational model of prefrontal control in free recall: Strategic memory use in the California Verbal Learning Task. *Journal of Cognitive Neuroscience*, *15*, 821–832. https://doi.org/10.1162/089892903322370744

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., & Wiener, S. I. (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal-prefrontal network upon learning. *Neuron*, *66*(6), 921–936. https://doi.org/10.1016/j.neuron.2010.05.013

Benjamin, A. S. (2007). Memory is more than just remembering: Strategic control of encoding, accessing memory, and making decisions. *Psychology of Learning and Motivation*, *48*, 175–223. https://doi.org/10.1016/s0079-7421(07)48005-7

Bentin, S., & Moscovitch, M. (1988). The time course of repetition effects for words and unfamiliar faces. *Journal of Experimental Psychology: General*, *117*(2), 148–160. https://doi.org/10.1037/0096-3445.117.2.148

Bhatarah, P., Ward, G., & Tan, L. (2006). Examining the relationship between free recall and immediate serial recall: The effect of concurrent task performance. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *32*(2), 215–229. https://doi.org/10.1037/0278-7393.32.2.215

Biggs, J. B. (1978). Individual and group differences in study processes. *British Journal of Educational Psychology*, *48*(3), 266–279. https://doi.org/10.1111/j.2044-8279.1978.tb03013.x

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*(2), 173–189. https://doi.org/10.1016/0010-0285(74)90009-7

Blumenfeld, R., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. *The Neuroscientist*, *13*(3), 280–291. https://doi.org/10.1177/1073858407299290

Bonnici, H. M., Chadwick, M. J., Lutti, A., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2012). Detecting representations of recent and remote autobiographical memories in vmPFC and hippocampus. *Journal of Neuroscience*, *32*(47), 16982–16991. https://doi.org/10.1523/jneurosci.2475-12.2012

Bouffard, N., Stokes, J., Kramer, H. J., & Ekstrom, A. D. (2018). Temporal encoding strategies result in boosts to final free recall performance comparable to spatial ones. *Memory & Cognition*, *46*(1), 17–31. https://doi.org/10.3758/s13421-017-0742-z

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, *49*(2), 229–240. https://doi.org/10.1080/00221309.1953.9710088

Bower, G. H., Monteiro, K. P., & Gilligan, S. G. (1978). Emotional mood as a context for learning and recall. *Journal of Learning and Verbal Behavior*, *17*, 573–585. https://doi.org/10.1016/s0022-5371(78)90348-1

Bradley, M. M., & Glenberg, A. M. (1983). Strengthening associations: Duration, attention, or relations? *Journal of Verbal Learning and Verbal Behavior*, *22*(6), 650–666. https://doi.org/10.1016/s0022-5371(83)90385-7

Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: Brain activity that predicts how well visual experience will be remembered. *Science*, *281*, 1185–1187. https://doi.org/10.1126/science.281.5380.1185

Buckner, R. L., & Wheeler, M. E. (2001). The cognitive neuroscience of remembering. *Nature Reviews Neuroscience*, *2*(9), 624–634. https://doi.org/10.1038/35090048

Butters, M. A., Kaszniak, A. W., Glisky, E. L., Eslinger, P. J., & Schacter, D. L. (1994). Recency discrimination deficits in frontal lobe patients. *Neuropsychology*, *8*(3), 343–353. https://doi.org/10.1037/0894-4105.8.3.343

Buzsáki, G., & Llinás, R. (2017). Space and time in the brain. *Science*, *358*(6362), 482–485. https://doi.org/10.1126/science.aan8869

Cabeza, R., Dolcos, F., Graham, R., & Nyberg, L. (2002). Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *NeuroImage*, *16*, 317–330. https://doi.org/10.1006/nimg.2002.1063

Camilleri, J. A., Müller, V. I., Fox, P., Laird, A. R., Hoffstaedter, F., Kalenscher, T., & Eickhoff, S. B. (2018). Definition and characterization of an extended multiple-demand network. *NeuroImage*, *165*, 138–147. https://doi.org/10.1016/j.neuroimage.2017.10.020

Campbell, K. L., & Hasher, L. (2018). Hyper-binding only apparent under fully implicit test conditions. *Psychology and Aging*, *33*(1), 176–181. https://doi.org/10.1037/pag0000216

Campbell, K. L., Trelle, A., & Hasher, L. (2014). Hyper-binding across time: Age differences in the effect of temporal proximity on paired-associate learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 293–299. https://doi.org/10.1037/a0034109

Caplan, J. B., Glaholt, M., & McIntosh, A. R. (2006). Linking associative and list memory: Pairs versus triples. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1244–1265. https://doi.org/10.1037/0278-7393.32.6.1244

Carroll, M., & Kirsner, K. (1982). Context and repetition effects in lexical decision and recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *21*(1), 55–69. https://doi.org/10.1016/s0022-5371(82)90445-5

Chan, S. C. Y., Applegate, M. C., Morton, N. W., Polyn, S. M., & Norman, K. A. (2017). Lingering representations of stimuli influence recall organization. *Neuropsychologia*, *97*, 72–82. https://doi.org/10.1016/j.neuropsychologia.2017.01.029

Clark, D. P. A., & Bruno, D. (2021). Time is of the essence: Exploring temporal and spatial organisation in episodic memory. *Quarterly Journal of Experimental Psychology*, *74*(8), 1406–1417. https://doi.org/10.1177/1747021821993823

Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia*, *6*(3), 255–265. https://doi.org/10.1016/0028-3932(68)90024-9

Cortis, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: Similar list length and output order effects for verbal and nonverbal stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1179–1214. https://doi.org/10.1037/xlm0000086

Craik, F. I. M. (2002). Levels of processing: Past, present... and future? *Memory*, *10*(5-6), 305–318. https://doi.org/10.1080/09658210244000135

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Curiel, J. M., & Radvansky, G. A. (1998). Mental organization of maps. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 202–214. https://doi.org/10.1037/0278-7393.24.1.202

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), 3–42. https://doi.org/10.1037/0033-295X.112.1.3

Davis, O. C., Geller, A. S., Rizzuto, D. S., & Kahana, M. J. (2008). Temporal associative processes revealed by intrusions in paired-associate recall. *Psychonomic Bulletin & Review*, *15*(1), 64–69. https://doi.org/10.3758/PBR.15.1.64

Delaney, P. F., & Knowles, M. E. (2005). Encoding strategy changes and spacing effects in the free recall of unmixed lists. *Journal of Memory and Language*, *52*(1), 120–130. https://doi.org/10.1016/j.jml.2004.09.002

Deuker, L., Bellmund, J. L., Navarro Schröder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *eLife*, *5*, e16534. https://doi.org/10.7554/eLife.16534

Dew, I. T. Z., Bayen, U. J., & Giovanello, K. S. (2007). Implicit relational memory in young and older adults. *Journal of Psychology*, *215*, 25–34. https://doi.org/10.1027/0044-3409.215.1.25

Dew, I. T. Z., & Cabeza, R. (2011). The porous boundaries between explicit and implicit memory: Behavioral and neural evidence. *Annals of the New York Academy of Sciences*, *1224*(1), 174–190. https://doi.org/10.1111/j.1749-6632.2010.05946.x

Dewhurst, S. A., Holmes, S. J., Brandt, K. R., & Dean, G. M. (2006). Measuring the speed of the conscious components of recognition memory: Remembering is faster than knowing. *Consciousness and Cognition*, *15*(1), 147–162. https://doi.org/10.1016/j.concog.2005.05.002

Diamond, N. B., & Levine, B. (2020). Linking detail to temporal structure in naturalistic-event recall. *Psychological Science*, *31*(12), 1557–1572. https://doi.org/10.1177/0956797620958651

Dimitrov, M., Granetz, J., Peterson, M., Hollnagel, C., Alexander, G., & Grafman, J. (1999). Associative learning impairments in patients with frontal lobe damage. *Brain and Cognition*, *41*(2), 213–230. https://doi.org/10.1006/brcg.1999.1121

Dobbins, I., Foley, H., Schacter, D., & Wagner, A. (2002). Executive control during episodic retrieval: Multiple prefrontal processes subserve source memory. *Neuron*, *35*(5), 989–996. https://doi.org/10.1016/S0896-6273(02)00858-9

Duarte, A., Ranganath, C., & Knight, R. T. (2005). Effects of unilateral prefrontal lesions on familiarity, recollection, and source memory. *Journal of Neuroscience*, *25*(36), 8333–8337. https://doi.org/10.1523/JNEUROSCI.1392-05.2005

DuBrow, S., Rouhani, N., Niv, Y., & Norman, K. A. (2017). Does mental context drift or shift? *Current Opinion in Behavioral Sciences*, *17*, 141–146. https://doi.org/10.1016/j.cobeha.2017.08.003

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Durgunoğlu, A. Y., & Neely, J. H. (1987). On obtaining episodic priming in a lexical decision task following paired-associate learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(2), 206–222. https://doi.org/10.1037/0278-7393.13.2.206

Eichenbaum, H., Yonelinas, A., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Eichenbaum, H. (2017a). Memory: Organization and control. *Annual Review of Psychology*, *68*, 19–45. https://doi.org/10.1146/annurev-psych-010416-044131

Eichenbaum, H. (2017b). Prefrontal-hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, *18*(9), 547–558. https://doi.org/10.1038/nrn.2017.74

Einstein, G. O., & Hunt, R. R. (1980). Levels of processing and organization: Additive effects of individual-item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(5), 588–598. https://doi.org/10.1037/0278-7393.6.5.588

Ekstrom, A. D., & Ranganath, C. (2018). Space, time, and episodic memory: The hippocampus is all over the cognitive map. *Hippocampus*, *28*(9), 680–687.

Engelkamp, J., & Zimmer, H. D. (1997). Sensory factors in memory for subject-performed tasks. *Acta Psychologica*, *96*(1), 43–60. https://doi.org/10.1016/s0001-6918(97)00005-x

Eysenck, M. W. (1979). Depth, elaboration, and distinctiveness. In F. I. M. Cermak L.S.; Craik (Ed.), *Levels of processing in human memory* (pp. 89–118). Psychology Press.

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*, *81*(5), 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*(2), 223–271. https://doi.org/10.1037/a0027371

Gabrieli, J. D. E. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology*, *49*, 87–115. https://doi.org/10.1146/annurev.psych.49.1.87

Gaines, J. J., Shapiro, A., Alt, M., & Benedict, R. H. B. (2006). Semantic clustering indexes for the Hopkins Verbal Learning Test-Revised: Initial exploration in elder control and dementia groups. *Applied Neuropsychology*, *13*, 213–222. https://doi.org/10.1207/s15324826an1304_2

Gershberg, F. B., & Shimamura, A. P. (1995). Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia*, *33*, 1305–1333. https://doi.org/10.1016/0028-3932(95)00103-a

Gibson, B. S., Healey, M. K., & Gondoli, D. M. (2019). ADHD reflects impaired externally directed and enhanced internally directed attention in the immediate free-recall task. *Journal of Abnormal Psychology*, *128*(2), 173–183. https://doi.org/10.1037/abn0000393

Gibson, B. S., Healey, M. K., Schor, D., & Gondoli, D. M. (2021). Space and time in the similarity structure of memory. *Psychonomic Bulletin & Review*, *28*, 2003–2011. https://doi.org/10.3758/s13423-021-01940-x

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 351–360. https://doi.org/10.1016/S0022-5371(66)80044-0

Glanzer, M. (1969). Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning and Verbal Behavior*, *8*, 105–111. https://doi.org/10.1016/s0022-5371(69)80018-6

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology*, *66*, 325–331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Goshen-Gottstein, Y., & Moscovitch, M. (1995). Repetition priming effects for newly formed associations are perceptually based: Evidence from shallow encoding and format specificity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1249–1262. https://doi.org/10.1037/0278-7393.21.5.1249

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 501–518. https://doi.org/10.1037/0278-7393.11.3.501

Graf, P., & Schacter, D. L. (1989). Unitization and grouping mediate dissociations in memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(5), 930–940. https://doi.org/10.1037/0278-7393.15.5.930

Grober, E., Lipton, R. B., Hall, C., & Crystal, H. (2000). Memory impairment on free and cued selective reminding predicts dementia. *Neurology*, *54*(4), 827–832. https://doi.org/10.1212/wnl.54.4.827

Hagen, J. W., Meacham, J. A., & Mesibov, G. (1970). Verbal labeling, rehearsal, and short-term memory. *Cognitive Psychology*, *1*(1), 47–58. https://doi.org/10.1016/0010-0285(70)90004-6

Hallock, H. L., Wang, A., & Griffin, A. L. (2016). Ventral midline thalamus is critical for hippocampal–prefrontal synchrony and spatial working memory. *Journal of Neuroscience*, *36*(32), 8372–8389. https://doi.org/10.1523/JNEUROSCI.0991-16.2016

Hasher, L., Lustig, C., & Zacks, R. T. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 227–249). Oxford University Press.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, *108*(3), 356–388. https://doi.org/10.1037/0096-3445.108.3.356

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, *22*, 193–225. https://doi.org/10.1016/s0079-7421(08)60041-9

Hazy, T. E., Frank, M. J., & OReilly, R. C. (2006). Banishing the homunculus: Making working memory work. *Neuroscience*, *139*(1), 105–118. https://doi.org/10.1016/j.neuroscience.2005.04.067

Healey, M. K., Long, N. M., & Kahana, M. J. (2019). Contiguity in episodic memory. *Psychonomic Bulletin & Review*, *26*(3), 699–720. https://doi.org/10.3758/s13423-018-1537-3

Healey, M. K., Mundorf, A. M. D., & Uitvlugt, M. G. (2020). Does depth of processing affect temporal contiguity? https://doi.org/10.31234/osf.io/dkxzj

Healey, M. K. (2018). Temporal contiguity in incidentally encoded memories. *Journal of Memory and Language*, *102*, 28–40. https://doi.org/10.1016/j.jml.2018.04.003

Healey, M. K., & Kahana, M. J. (2014). Is memory search governed by universal principles or idiosyncratic strategies? *Journal of Experimental Psychology: General*, *143*(2), 575–596. https://doi.org/10.1037/a0033715

Healey, M. K., & Kahana, M. J. (2016). A four-component model of age-related memory change.

Psychological Review, *123*(1), 23–69. https://doi.org/10.1037/rev0000015

Healey, M. K., & Kahana, M. J. (2020). Age-related differences in the temporal dynamics of spectral power during memory encoding. *PLOS ONE*, *15*(1), e0227274. https://doi.org/10.1371/journal.pone.0227274

Healey, M. K., Ngo, K. W. J., & Hasher, L. (2014). Below–baseline suppression of competitors during interference resolution by younger but not older adults. *Psychological Science*, *25*(1), 145–151. https://doi.org/10.1177/0956797613501169

Healey, M. K., & Uitvlugt, M. G. (2019). The role of control processes in temporal and semantic contiguity. *Memory & Cognition*, *47*(4), 719–737. https://doi.org/10.3758/s13421-019-00895-8

Healey, M. K., & Wahlheim, C. N. (2023). PEPPR: A post-encoding pre-production reinstatement model of dual-list free recall. *Memory & Cognition*. https://doi.org/10.3758/s13421-023-01453-z

Hintzman, D. L. (2016). Is memory organized by temporal contiguity? *Memory & Cognition*, *44*(3), 365–375. https://doi.org/10.3758/s13421-015-0573-8

Hintzman, D. L., Block, R. A., & Inskeep, N. R. (1972). Memory for mode of input. *Journal of Verbal Learning and Verbal Behavior*, *11*, 741–749. https://doi.org/10.1016/s0022-5371(72)80008-2

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science*, *6*(3), 253–271. https://doi.org/10.1177/1745691611406924

Hirshman, E., & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 484–494. https://doi.org/10.1037/0278-7393.14.3.484

Hirst, W., & Volpe, B. T. (1988). Memory strategies with brain damage. *Brain and Cognition*, *8*(3), 379–408. https://doi.org/10.1016/0278-2626(88)90060-7

Hong, B., Pace-Tonna, C. A., Barense, M. D., & Mack, M. L. (2022). Emphasizing temporal and semantic associations from encoding affects free recall at retrieval [preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/q4dpg

Hong, M. K., Fazio, L. K., & Polyn, S. M. (2019). Examining factors that eliminate contiguity in free recall. *Poster presented at Psychonomic Society, Montréal, Québec, Canada*.

Hong, M. K., Gunn, J. B., Fazio, L. K., & Polyn, S. M. (2023). The modulation and elimination of temporal organization in free recall. *Journal of Experimental Psychology: Learning,* *Memory, and Cognition*.

Hong, M. K., Polyn, S. M., & Fazio, L. K. (2019). Examining the episodic context account: Does retrieval practice enhance memory for context? *Cognitive Research: Principles and Implications*, *4*(1), 46. https://doi.org/10.1186/s41235-019-0202-3

Horwath, E. A., Rouhani, N., DuBrow, S., & Murty, V. P. (2023). Value restructures the organization of free recall. *Cognition*, *231*, 105315. https://doi.org/10.1016/j.cognition.2022.105315

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, *112*(1), 75–116. https://doi.org/10.1037/0033-295x.112.1.75

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. https://doi.org/10.1037/0278-7393.25.4.923

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Howard, M. W., Shankar, K. H., Aue, W. R., & Criss, A. H. (2015). A distributed representation of internal time. *Psychological Review*, *122*(1), 24–53. https://doi.org/10.1037/a0037840

Howard, M. W., Viskontas, I. V., Shankar, K. H., & Fried, I. (2012). Ensembles of human MTL neurons "jump back in time" in response to a repeated stimulus. *Hippocampus*, *22*(9), 1833–1847. https://doi.org/10.1002/hipo.22018

Howard, M. W., Kahana, M. J., & Wingfield, A. (2006). Aging and contextual binding: Modeling recency and lag-recency effects with the temporal context model. *Psychonomic Bulletin & Review*, *13*(3), 439–445. https://doi.org/10.3758/bf03193867

Hunt, R. R., Smith, R. E., & Dunlap, K. R. (2011). How does distinctive processing reduce false recall? *Journal of Memory and Language*, *65*(4), 378–389. https://doi.org/10.1016/j.jml.2011.06.003

Hyde, T. S., & Jenkins, J. J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, *82*(3), 472–481. https://doi.org/10.1037/h0028372

Hyman, J., Zilli, E., Paley, A., & Hasselmo, M. (2010). Working memory performance correlates with prefrontal-hippocampal theta interactions but not with prefrontal neuron firing rates. *Frontiers in Integrative Neuroscience*, *4*, 1–13. https://doi.org/10.3389/neuro.07.002.2010

İmamoğlu, A., Foubert, C., Healey, M. K., Langella, S., Belger, A., Giovanello, K. S., &

Wahlheim, C. N. (2022). Episodic memory impairment in children and adolescents at risk for schizophrenia: A role for context processing. *Schizophrenia Research: Cognition*, *28*, 100241. https://doi.org/10.1016/j.scog.2022.100241

Incisa della Rocchetta, A. (1986). Classification and recall of pictures after unilateral frontal or temporal lobectomy. *Cortex*, *22*(2), 189–211. https://doi.org/10.1016/S0010-9452(86)80045-4

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541. https://doi.org/10.1016/0749-596x(91)90025-f

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, *27*(8), 1043–1056. https://doi.org/10.1016/0028-3932(89)90184-x

Jayakumar, M., Balusu, C., & Aly, M. (2023). Attentional fluctuations and the temporal organization of memory. *Cognition*, *235*, 105408. https://doi.org/10.1016/j.cognition.2023.105408

Jenkins, L. J., & Ranganath, C. (2010). Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *Journal of Neuroscience*, *30*(46), 15558–15565. https://doi.org/10.1523/jneurosci.1337-10.2010

Jennings, J. M., & Jacoby, L. L. (1993). Automatic versus intentional uses of memory: Aging, attention, and control. *Psychology and Aging*, *8*(2), 283. https://doi.org/10.1037/0882-7974.8.2.283

Jetter, W., Poser, U., Freeman, R. B., & Markowitsch, H. J. (1986). A verbal long term memory deficit in frontal lobe damaged patients. *Cortex*, *22*(2), 229–242. https://doi.org/10.1016/s0010-9452(86)80047-8

Jones, M. W., & Wilson, M. A. (2005). Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. *PLOS Biology*, *3*(12), e402. https://doi.org/10.1371/journal.pbio.0030402

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109. https://doi.org/10.3758/BF03197276

Kahana, M. J., & Howard, M. W. (2005). Spacing and lag effects in free recall of pure lists. *Psychonomic Bulletin & Review*, *12*(1), 159–164. https://doi.org/10.3758/bf03196362

Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 530–540. https://doi.org/10.1037/0278-7393.28.3.530

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs.

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. *Psychology of Learning and Motivation*, *61*, 237–284. https://doi.org/https://doi.org/10.1016/B978-0-12-800283-4.00007-1

Kim, H., Daselaar, S. M., & Cabeza, R. (2010). Overlapping brain activity between episodic memory encoding and retrieval: Roles of the task-positive and task-negative networks. *NeuroImage*, *49*(1), 1045–1054. https://doi.org/10.1016/j.neuroimage.2009.07.058

Koriat, A., & Melkman, R. (1987). Depth of processing and memory organization. *Psychological Research*, *49*(2-3), 173–181. https://doi.org/10.1007/BF00308684

Kowialiewski, B., Krasnoff, J., Mizrak, E., & Oberauer, K. (2022). The semantic relatedness effect in serial recall: Deconfounding encoding and recall order. *Journal of Memory and Language*, *127*, 104377. https://doi.org/10.1016/j.jml.2022.104377

Kragel, J. E., & Voss, J. L. (2021). Temporal context guides visual exploration during scene recognition. *Journal of Experimental Psychology: General*, *150*(5), 873–889. https://doi.org/10.1037/xge0000827

Kuhl, B. A., & Wagner, A. D. (2009). Strategic control of memory. In *Encyclopedia of neuroscience* (pp. 437–444). Elsevier Ltd. https://doi.org/10.1016/b978-008045046-9.00424-1

Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology*, *34*(5-6), 419–426. https://doi.org/10.1080/002075999399774

Lazarus, L. T., Mundorf, A. M. D., Uitvlugt, M. G., & Healey, M. K. (in prep). An item-order model of reduced temporal contiguity for distinct items. *Manuscript submitted for publication*.

Lehman, M., & Malmberg, K. J. (2013). A buffer model of memory encoding and temporal correlations in retrieval. *Psychological Review*, *120*(1), 155–189. https://doi.org/10.1037/a0030851

Lehmann, M., & Hasselhorn, M. (2010). The dynamics of free recall and their relation to rehearsal between 8 and 10 years of age. *Child Development*, *81*(3), 1006–1020. https://doi.org/10.1111/j.1467-8624.2010.01448.x

Lehmann, M., & Hasselhorn, M. (2012). Rehearsal dynamics in elementary school children. *Journal of Experimental Child Psychology*, *111*(3), 552–560. https://doi.org/10.1016/j.jecp.2011.10.013

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015). Expanding the scope of memory search: Intralist and interlist effects in free recall. *Psychological Review*, *122*(2), 337–363. https://doi.org/10.1037/a0039036

Long, N. M., Danoff, M. S., & Kahana, M. J. (2015). Recall dynamics reveal the retrieval of emotional context. *Psychonomic Bulletin & Review*, *22*(5), 1328–1333. https://doi.org/10.3758/s13423-014-0791-2

Long, N. M., & Kahana, M. J. (2017). Modulation of task demands suggests that semantic processing interferes with the formation of episodic associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 167–176. https://doi.org/10.1037/xlm0000300

Mack, C. C., Cinel, C., Davies, N., Harding, M., & Ward, G. (2017). Serial position, output order, and list length effects for words presented on smartphones over very long intervals. *Journal of Memory and Language*, *97*, 61–80. https://doi.org/10.1016/j.jml.2017.07.009

Malek-Ahmadi, M., Raj, A., & Small, B. J. (2011). Semantic clustering as a neuropsychological predictor for amnestic-MCI. *Aging, Neuropsychology, and Cognition*, *18*(3), 280–292. https://doi.org/10.1080/13825585.2010.540642

Mandler, G. (1967). Organization and memory. *Psychology of Learning and Motivation*, *1*, 327–372. https://doi.org/10.1016/b978-0-12-121050-2.50014-4

Mangels, J. (1997). Strategic processing and memory for temporal order in patients with frontal lobe lesions. *Neuropsychology*, *11*(2), 207–221. https://doi.org/10.1037/0894-4105.11.2.207

Manning, J. R., Polyn, S. M., Baltuch, G., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences*, *108*(31), 12893–12897. https://doi.org/10.1073/pnas.1015174108

Manns, J. R., Howard, M. W., & Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, *56*(3), 530–540. https://doi.org/10.1016/j.neuron.2007.08.017

Martin, A., Brouwers, P., Cox, C., & Fedio, P. (1985). On the nature of the verbal memory deficit in Alzheimer's disease. *Brain and Language*, *25*(2), 323–341. https://doi.org/10.1016/0093-934x(85)90088-4

Mastrogiuseppe, M., Bertelsen, N., Bedeschi, M. F., & Lee, S. A. (2019). The spatiotemporal organization of episodic memory and its disruption in a neurodevelopmental disorder. *Scientific Reports*, *9*(1), 18447. https://doi.org/10.1038/s41598-019-53823-w

Mazuryk, G. F., & Lockhart, R. S. (1974). Negative recency and levels of processing in free recall. *Canadian Journal of Psychology*, *28*(1), 114–123. https://doi.org/10.1037/h0081971

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295x.102.3.419

McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review*, *15*(2), 237–255. https://doi.org/10.3758/pbr.15.2.237

McDaniel, M. A., Cahill, M., Bugg, J. M., & Meadow, N. G. (2011). Dissociative effects of orthographic distinctiveness in pure and mixed lists: An item-order account. *Memory & Cognition*, *39*(7), 1162–1173. https://doi.org/10.3758/s13421-011-0097-9

McKoon, G., & Ratcliff, R. (1979). Priming in episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *18*(4), 463–480. https://doi.org/10.1016/s0022-5371(79)90255-x

McKoon, G., & Ratcliff, R. (1986). Automatic activation of episodic information in a semantic memory task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 108. https://doi.org/10.1037/0278-7393.12.1.108

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miller, J. F., Neufang, M., Solway, A., Brandt, A., Trippel, M., Mader, I., Hefft, S., Merkow, M., Polyn, S. M., Jacobs, J., Kahana, M. J., & Schulze-Bonhage, A. (2013). Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science*, *342*(6162), 1111–1114. https://doi.org/10.1126/science.1244056

Milner, B., Petrides, M., & Smith, M. L. (1985). Frontal lobes and the temporal organization of memory. *Human Neurobiology*, *4*(3), 137–142. http://europepmc.org/abstract/MED/4066424

Moreton, B. J., & Ward, G. (2010). Time scale similarity and long-term memory for autobiographical events. *Psychonomic Bulletin & Review*, *17*(4), 510–515. https://doi.org/10.3758/pbr.17.4.510

Morton, N. W., & Polyn, S. M. (2016). A predictive framework for evaluating models of semantic organization in free recall. *Journal of Memory and Language*, *86*, 119–140. https://doi.org/10.1016/j.jml.2015.10.002

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, *4*(3), 257–267. https://doi.org/10.1162/jocn.1992.4.3.257

Moscovitch, M., & Craik, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior*, *15*(4), 447–458. https://doi.org/10.1016/s0022-5371(76)90040-2

Moscovitch, M., & Winocur, G. (1995). Frontal lobes, memory, and aging. *Annals of New York Academy of Sciences*, *769*, 119–150. https://doi.org/10.1111/j.1749-6632.1995.tb38135.x

Mundorf, A. M. D., Uitvlugt, M. G., & Healey, M. K. (2023). Incidentally encoded temporal associations produce priming in implicit memory. *Psychonomic Bulletin & Review*. https://doi.org/https://doi.org/10.3758/s13423-023-02351-w

Mundorf, A. M. D., Lazarus, L. T. T., Uitvlugt, M. G., & Healey, M. K. (2021). A test of retrieved context theory: Dynamics of recall after incidental encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(8), 1264–1287. https://doi.org/10.1037/xlm0001001

Mundorf, A. M. D., Uitvlugt, M. G., & Healey, M. K. (2022). Does depth of processing affect temporal contiguity? *Psychonomic Bulletin & Review*, *29*, 2229–2239. https://doi.org/10.3758/s13423-022-02112-1

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482–488. https://doi.org/10.1037/h0045106

Murdock, B. B. (1972). Short-term memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory.* (pp. 67–127, Vol. 5). Academic Press.

Murdock, B. B. (1974). *Human memory: Theory and data*. Lawrence Erlbaum; Associates.

Murphy, D. H., & Castel, A. D. (2021). The role of attention and ageing in the retrieval dynamics of value-directed remembering. *Quarterly Journal of Experimental Psychology*, *75*(5), 954–968. https://doi.org/10.1177/17470218211046612

Murre, J. M. J. (2021). The Godden and Baddeley (1975) experiment on context-dependent memory on land and underwater: A replication. *Royal Society Open Science*, *8*(11). https://doi.org/10.1098/rsos.200724

Murty, V. P., McKinney, R. A., DuBrow, S., Jalbrzikowski, M., Haas, G. L., & Luna, B. (2018). Differential patterns of contextual organization of memory in first-episode psychosis. *Schizophrenia*, *4*(1). https://doi.org/10.1038/s41537-018-0046-8

Nairne, J. S., Cogdill, M., & Lehman, M. (2017). Adaptive memory: Temporal, semantic, and rating-based clustering following survival processing. *Journal of Memory and Language*, *93*, 304–314. https://doi.org/10.1016/j.jml.2016.10.009

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 702–709. https://doi.org/10.1037/0278-7393.17.4.702

Navawongse, R., & Eichenbaum, H. (2013). Distinct pathways for rule-based retrieval and spatial mapping of memory representations in hippocampal neurons. *Journal of Neuroscience*, *33*(3), 3891–3912. https://doi.org/10.1523/jneurosci.3891-12.2013

Naveh-Benjamin, M. (2000). Adult-age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1170–1187. https://doi.org/10.1037/0278-7393.26.5.1170

Neath, I. (1993). Contextual and distinctive processes and the serial position function. *Journal of Memory and Language*, *32*(6), 820–840. https://doi.org/10.1006/jmla.1993.1041

Nguyen, K., & McDaniel, M. A. (2015). The picture complexity effect: Another list composition paradox. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1026. https://doi.org/10.1037/xlm0000071

Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S., & Sederberg, P. B. (2015). Human hippocampus represents space and time during retrieval of real-world memories. *Proceedings of the National Academy of Sciences*, *112*(35), 11078–11083. https://doi.org/10.1073/pnas.1507104112

Nilsson, L.-G. (1974). Further evidence for organization by modality in free recall. *Journal of Experimental Psychology*, *103*(5), 948–957. https://doi.org/10.1037/h0037406

Old, S., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*(1), 104–118.

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* (pp. 375–411). Cambridge University Press.

O'Reilly, R. C., Herd, S. A., & Pauli, W. M. (2010). Computational models of cognitive control. *Current Opinion in Neurobiology*, *20*(2), 257–261. https://doi.org/10.1016/j.conb.2010.01.008

Osth, A., & Fox, J. (2019). Are associations formed across word pairs? A test of learning by temporal contiguity in associative recognition. *Psychonomic Bulletin & Review*, *26*, 1650–1656. https://doi.org/10.3758/s13423-019-01616-7

Pacheco, D., & Verschure, P. F. M. J. (2018). Long-term spatial clustering in free recall. *Memory*, *26*(6), 798–806. https://doi.org/https://doi.org/10.1177/1747021821993823

Pajkossy, P., Keresztes, A., & Racsmány, M. (2017). The interplay of trait worry and trait anxiety in determining episodic retrieval: The role of cognitive control. *The Quarterly Journal of Experimental Psychology*, *70*(11), 2234–2250. https://doi.org/10.1080/17470218.2016.1230142

Palombo, D. J., Lascio, J. M. D., Howard, M. W., & Verfaellie, M. (2019). Medial temporal lobe amnesia is associated with a deficit in recovering temporal context. *Journal of Cognitive Neuroscience*, *31*(2), 236–248. https://doi.org/10.1162/jocn_a_01344

Pathman, T., Deker, L., Parmar, P. K., Adkins, M. C., & Polyn, S. M. (2023). Children's memory "in the wild": Examining the temporal organization of free recall from a week-long camp at a local zoo. *Cognitive Research: Principles and Implications*, *8*(1), 6. https://doi.org/10.1186/s41235-022-00452-z

Picton, T. W., Stuss, D. T., Alexander, M. P., Shallice, T., Binns, M. A., & Gillingham, S. (2007). Effects of focal frontal lesions on response inhibition. *Cerebral Cortex*, *17*(4), 826–838. https://doi.org/https://doi.org/10.1093/cercor/bhk031

Place, R., Farovik, A., Brockmann, M., & Eichenbaum, H. (2016). Bidirectional prefrontal-hippocampal interactions support context-guided memory. *Nature Neuroscience*, *19*(8), 992–994. https://doi.org/10.1038/nn.4327

Poldrack, R. A., & Cohen, N. J. (1997). Priming of new associations in reading time: What is learned? *Psychonomic Bulletin & Review*, *4*(3), 398–402. https://doi.org/10.3758/bf03210800

Polyn, S. M., Erlikhman, G., & Kahana, M. J. (2011). Semantic cuing and the scale-insensitivity of recency and contiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 766–775. https://doi.org/10.1037/a0022475

Polyn, S. M., & Kahana, M. J. (2008). Memory search and the neural representation of context. *Trends in Cognitive Sciences*, *12*(1), 24–30. https://doi.org/10.1016/j.tics.2007.10.010

Polyn, S. M., Kragel, J. E., Morton, N. W., McCluey, J. D., & Cohen, Z. D. (2012). The neural dynamics of task context in free recall. *Neuropsychologia*, *50*(4), 447–457. https://doi.org/10.1016/j.neuropsychologia.2011.08.025

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009a). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129–156. https://doi.org/10.1037/a0014420

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009b). Task context and organization in free recall. *Neuropsychologia*, *47*, 2158–2163. https://doi.org/10.1016/j.neuropsychologia.2009.02.013

Polyn, S. M., McCluey, J. D., Morton, N. W., Woolard, A. A., Luksik, A. S., & Heckers, S. (2015). Temporal context and the organisational impairment of memory search in schizophrenia. *Cognitive Neuropsychiatry*, *20*(4), 296–310. https://doi.org/10.1080/13546805.2015.1031372

Postman, L. (1971). Organization and interference. *Psychological Review*, *78*(4), 290–302. https://doi.org/10.1037/h0031031

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, *23*(17), R764–R773. https://doi.org/10.1016/j.cub.2013.05.041

Pu, Y., Kong, X.-Z., Ranganath, C., & Melloni, L. (2022). Event boundaries shape temporal organization of memory by resetting temporal context. *Nature Communications*, *13*(1), 622. https://doi.org/10.1038/s41467-022-28216-9

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 207–262, Vol. 14). Academic Press.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*(2), 93–134. https://doi.org/10.1037/0033-295X.88.2.93

Raaijmakers, J. G. (2005). Modeling implicit and explicit memory. *Human learning and memory: Advances in theory and application: The 4th Tsukuba International Conference on Memory*, 85–105.

Ranganath, C., Cohen, M. X., Dam, C., & D'Esposito, M. (2004). Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *Journal of Neuroscience*, *24*(16), 3917–3925. https://doi.org/10.1523/jneurosci.5053-03.2004

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*(3), 385–408. https://doi.org/10.1037/0033-295x.95.3.385

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*(3), 446–461. https://doi.org/10.1037/0033-2909.86.3.446

Raz, N., Gunning, F. M., Head, D., Dupuis, J. H., McQuain, J., Briggs, S. D., Loken, W. J., Thornton, A. E., & Acker, J. D. (1997). Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*, *7*(3), 268–282. https://doi.org/10.1093/cercor/7.3.268

Reuter-Lorenz, P. A., & Park, D. C. (2010). Human neuroscience and the aging mind: A new look at old problems. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65*(4), 405–415. https://doi.org/10.1093/geronb/gbq035

Rich, E. L., & Shapiro, M. (2009). Rat prefrontal cortical neurons selectively code strategy switches. *Journal of Neuroscience*, *29*(22), 7208–7219. https://doi.org/10.1523/jneurosci.6068-08.2009

Robin, J., Wynn, J., & Moscovitch, M. (2016). The spatial scaffold: The effects of spatial context on memory for events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(2), 308. https://doi.org/10.1037/xlm0000167

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin*, *76*(1), 45–48.

Rogerson, T., Cai, D. J., Frank, A., Sano, Y., Shobe, J., Lopez-Aranda, M. F., & Silva, A. J. (2014). Synaptic tagging during memory allocation. *Nature Reviews Neuroscience*, *15*(3), 157–169. https://doi.org/10.1038/nrn3667

Rose, N. S., & Craik, F. I. M. (2012). A processing approach to the working memory/long-term memory distinction: Evidence from the levels-of-processing span task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1019–1029. https://doi.org/10.1037/a0026976

Roux, F., Armstrong, B. C., & Carreiras, M. (2017). Chronset: An automated tool for detecting speech onset. *Behavior Research Methods*, *49*(5), 1864–1881. https://doi.org/10.3758/s13428-016-0830-1

Sadeh, T., Moran, R., & Goshen-Gottstein, Y. (2015). When items 'pop into mind': Variability in temporal-context reinstatement in free-recall. *Psychonomic Bulletin & Review*, *22*(3), 779–790. https://doi.org/10.3758/s13423-014-0746-7

Sahakyan, L., & Smith, J. R. (2014). A long time ago, in a context far, far away: Retrospective time estimates and internal context change. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 86–93. https://doi.org/10.1037/a0034250

Sahakyan, L., & Kwapil, T. R. (2018). Moving beyond summary scores: Decomposing free recall performance to understand episodic memory deficits in schizotypy. *Journal of Experimental Psychology: General*, *147*(12), 1919–1930. https://doi.org/10.1037/xge0000401

Schwartz, G., Howard, M. W., Jing, B., & Kahana, M. J. (2005). Shadows of the past: Temporal retrieval effects in recognition memory. *Psychological Science*, *16*(11), 898–904. https://doi.org/10.1111/j.1467-9280.2005.01634.x

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*, 11–21. https://doi.org/10.1136/jnnp.20.1.11

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*(4), 893–812. https://doi.org/10.1037/a0013396

Sederberg, P. B., Miller, J. F., Howard, M. W., & Kahana, M. J. (2010). The temporal contiguity effect predicts episodic memory performance. *Memory & Cognition*, *38*(6), 689–699. https://doi.org/10.3758/MC.38.6.689

Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., Litt, B., Brandt, A., & Kahana, M. J. (2007). Gamma oscillations distinguish true from false memories. *Psychological Science*, *18*(11), 927–932. https://doi.org/10.1111/j.1467-9280.2007.02003.x

Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., McCarthy, D. C., Brandt, A., Tully, M. S., & Kahana, M. J. (2007). Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cerebral Cortex*, *17*(5), 1190–1196. https://doi.org/10.1093/cercor/bhl030

Sederberg, P. B., Gershman, S. J., Polyn, S. M., & Norman, K. A. (2011). Human memory reconsolidation can be explained using the temporal context model. *Psychonomic Bulletin & Review*, *18*(3), 455–468. https://doi.org/10.3758/s13423-011-0086-9

Seiver, J. G., Pires, M., Awan, F., & Thompson, W. (2019). Retention of word pairs as a function of level of processing, instruction to remember, and delay. *Journal of Cognitive Psychology*, *31*(7), 665–682. https://doi.org/10.1080/20445911.2019.1664555

Shin, Y. S., Masís-Obando, R., Keshavarzian, N., Dáve, R., & Norman, K. A. (2021). Context-dependent memory effects in two immersive virtual reality environments: On Mars and underwater. *Psychonomic Bulletin & Review*, *28*(2), 574–582. https://doi.org/https://doi.org/10.3758/s13423-020-01835-3

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Siddiqui, A. P., & Unsworth, N. (2011). Investigating the role of emotion during the search process in free recall. *Memory & Cognition*, *39*(8), 1387–1400. https://doi.org/10.3758/s13421-011-0125-9

Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *40*(3), 755–764. https://doi.org/10.1037/a0035585

Silva, A. J., Zhou, Y., Rogerson, T., Shobe, J., & Balaji, J. (2009). Molecular and cellular approaches to memory allocation in neural circuits. *Science*, *326*(5951), 391–395. https://doi.org/10.1126/science.1174519

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, *4*(8), 637–648. https://doi.org/10.1038/nrn1178

Smith, M. C., Macleod, C. M., Bain, J. D., & Hoppe, R. B. (1989). Lexical decision as an indirect test of memory: Repetition priming and list-wide priming as a function of type of encoding.

Journal of Experimental Psychology: Learning, Memory and Cognition, *15*(6), 1109–1118. https://doi.org/10.1037/0278-7393.15.6.1109

Smith, T. A., Hasinski, A. E., & Sederberg, P. B. (2013). The context repetition effect: Predicted events are remembered better, even when they don't happen. *Journal of Experimental Psychology: General*, *142*(4), 1298–1308. https://doi.org/10.1037/a0034067

Spieler, D. H., & Balota, D. A. (1996). Characteristics of associative learning in younger and older adults: Evidence from an episodic priming paradigm. *Psychology and Aging*, *11*(4), 607–620. https://doi.org/10.1037/0882-7974.11.4.607

Spillers, G. J., & Unsworth, N. (2011). Variation in working memory capacity and temporal–contextual retrieval from episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1532–1539. https://doi.org/10.1037/a0024852

Squire, L. R., Clark, R. E., & Knowlton, B. J. (2001). Retrograde amnesia. *Hippocampus*, *11*(1), 50–55. https://doi.org/10.1002/1098-1063(2001)11:1<50::aid-hipo1019>3.0.co;2-g

Staresina, B. P., & Davachi, L. (2006). Differential encoding mechanisms for subsequent associative recognition and free recall. *Journal of Neuroscience*, *26*(36), 9162–9172. https://doi.org/10.1523/jneurosci.2877-06.2006

Staresina, B. P., & Davachi, L. (2009). Mind the gap: Binding experiences across space and time in the human hippocampus. *Neuron*, *63*(2), 267–276. https://doi.org/10.1016/j.neuron.2009.06.024

Stefanidi, A., & Brewer, G. A. (2015). *The recall dynamics of importance in delayed free recall* [Master's thesis, Arizona State University].

Stefanidi, A., Ellis, D. M., & Brewer, G. A. (2018). Free recall dynamics in value-directed remembering. *Journal of Memory and Language*, *100*, 18–31. https://doi.org/10.1016/j.jml.2017.11.004

Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In A. F. Healy (Ed.), *Cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer*. American Psychological Association. https://doi.org/10.1037/10895-018

Strange, B. A., Otten, L. J., Josephs, O., Rugg, M. D., & Dolan, R. J. (2002). Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. *Journal of Neuroscience*, *22*(2), 523–528. https://doi.org/10.1523/jneurosci.22-02-00523.2002

Stuss, D. T., Alexander, M. P., Palumbo, C. L., Buckle, L., Sayer, L., & Pogue, J. (1994). Organizational strategies with unilateral or bilateral frontal lobe injury in word learning tasks. *Neuropsychology*, *8*(3), 355. https://doi.org/10.1037/0894-4105.8.3.355

Summerfield, C. (2006). *Cognitive control during episodic memory formation* [Doctoral dissertation]. Columbia University.

Summerfield, J. J., Lepsien, J., Gitelman, D. R., Mesulam, M. M., & Nobre, A. C. (2006). Orienting attention based on long-term memory experimence. *Neuron*, *49*(6), 905–916. https://doi.org/10.1016/j.neuron.2006.01.021

Szczepanski, S. M., & Knight, R. T. (2014). Insights into human behavior from lesions to the prefrontal cortex. *Neuron*, *83*(5), 1002–1018. https://doi.org/10.1016/j.neuron.2014.08.011

Talamonti, D., Koscik, R., Johnson, S., & Bruno, D. (2021). Temporal contiguity and ageing: The role of memory organization in cognitive decline. *Journal of Neuropsychology*, *15*, 53–65. https://doi.org/10.1111/jnp.12219

Talamonti, D., Montgomery, C. A., Clark, D. P. A., & Bruno, D. (2020). Age-related prefrontal cortex activation in associative memory: An fNIRS pilot study. *NeuroImage*, *222*, 117223. https://doi.org/10.1016/j.neuroimage.2020.117223

Talmi, D., Lohnas, L. J., & Daw, N. D. (2019). A retrieved context model of the emotional modulation of memory. *Psychological Review*, *126*(4), 455–485. https://doi.org/10.1037/rev0000132

Toro-Serey, C., Bright, I. M., Wyble, B., & Howard, M. (2019). Rapid presentation rate negatively impacts the contiguity effect in free recall [preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/qb5sx

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory.* (pp. 381–403). Academic Press.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 381–391. https://doi.org/10.1016/S0022-5371(66)80048-8

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review*, *69*(4), 344–354. https://doi.org/10.1037/h0043150

Uitvlugt, M. G., & Healey, M. K. (2019). Temporal proximity links unrelated news events in memory. *Psychological Science*, *30*(1), 92–104. https://doi.org/10.1177/0956797618808474

Uitvlugt, M. G., & Healey, M. K. (2020). The temporal contiguity effect is negatively correlated with episodic memory performance and fluid intelligence. *Poster presented at the Con-*

text and Episodic Memory Symposium. https://cbcc.psy.msu.edu/presentations/Uitvlugt_CEMS_2020.pdf

Unsworth, N. (2008). Exploring the retrieval dynamics of delayed and final free recall: Further evidence for temporal-contextual search. *Journal of Memory and Language*, *59*(2), 223–236. https://doi.org/10.1016/j.jml.2008.04.002

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104–32. https://doi.org/10.1037/0033-295X.114.1.104

Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(1), 50–61. https://doi.org/10.1037/xlm0000148

Unsworth, N., Miller, A. L., & Robison, M. K. (2019). Individual differences in encoding strategies and free recall dynamics. *Quarterly Journal of Experimental Psychology*, *72*(10), 2495–2508. https://doi.org/10.1177/1747021819847441

Viskontas, I., Carr, V. A., Engel, S., & Knowlton, B. (2009). The neural correlates of recollection: hippocampal activation declines as episodic memory fades. *Hippocampus*, *19*(3), 265–272. https://doi.org/10.1002/hipo.20503

Wagner, A. (2002). Cognitive control and episodic memory: Contributions from the prefrontal cortex. In L. R. Squire & D. L. Schacter (Eds.), *Neuropsychology of memory* (3rd ed., pp. 174–192). The Guilford Press.

Wahlheim, C. N., & Huff, M. J. (2015). Age differences in the focus of retrieval: Evidence from dual-list free recall. *Psychology and Aging*, *30*(4), 768–780. https://doi.org/10.1037/pag0000049

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. https://doi.org/10.1037/a0020122

West, R. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*, 272–292. https://doi.org/10.1037/0033-2909.120.2.272

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1995). Frontal lobe damage produces episodic memory impairment. *Journal of the International Neuropsychological Society*, *1*, 525–536. https://doi.org/10.1017/s1355617700000655

Whiffen, J. W., & Karpicke, J. D. (2017). The role of episodic context in retrieval practice effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(7), 1036–1046. https://doi.org/10.1037/xlm0000379

White, I. M., & Wise, S. P. (1999). Rule-dependent neuronal activity in the prefrontal cortex. *Experimental Brain Research*, *126*, 315–335. https://doi.org/10.1007/s002210050740

Wise, S. P., Murray, E. A., & Gerfen, C. R. (1996). The frontal cortex-basal ganglia system in primates. *Critical Reviews in Neurobiology*, *10*(3-4), 317–356. https://doi.org/10.1615/critrevneurobiol.v10.i3-4.30

Yeung, N., Nystrom, L. E., Aronson, J. A., & Cohen, J. D. (2006). Between-task competition and cognitive control in task switching. *Journal of Neuroscience*, *26*(5), 1429–1438. https://doi.org/10.1523/jneurosci.3109-05.2006

Yonelinas, A. P., Kroll, N. E., Quamme, J. R., Lazzara, M. M., Sauvé, M. J., Widaman, K. F., & Knight, R. T. (2002). Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nature Neuroscience*, *5*(11), 1236–1241. https://doi.org/10.1038/nn961

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, *25*(11), 3002–3008. https://doi.org/10.1523/JNEUROSCI.5295-04.2005

Zeelenberg, R., Pecher, D., & Raaijmakers, J. G. W. (2003). Associative repetition priming: A selective review and theoretical implications. In J. S. Bowers & C. S. Marsolek (Eds.), *Rethinking implicit memory* (pp. 261–283). Oxford University Press.

Zhang, Q., Griffiths, T. L., & Norman, K. A. (2022). Optimal policies for free recall. *Psychological Review*, *130*(4), 1104–1124. https://doi.org/10.1037/rev0000375

APPENDIX

There are many versions of retrieved context models tailored for different purposes. The retrieved context model implemented in Chapter 4 is a variation of the Context Maintenance and Retrieval (CMR) model proposed by Polyn et al. (2009a) which follows the semantic context version of CMR described by Morton and Polyn (2016) except where noted. The model implementation used in Chapter 5 is based on the Post-Encoding Pre-Production Reinstatement (PEPPR) model introduced by Healey and Wahlheim (2023) which has been modified to include representations of task goals, task-relevant features (in this case, the three locations and three dishes associated with study items), and the relationship between them. Because CMR and PEPPR overlap in many respects, the two models used in Chapters 4 and 5 are described together, and any differences are noted.

Associative Matrices

The model contains two representational layers, the feature layer (F), where currently active items are represented in high-dimensional space, and the context layer (C), where the state of mental context is represented in a similar high-dimensional space. Activation of specific items or contexts within each of these representational spaces is defined as a vector, f or c respectively. These vectors have one node for each list item, one node representing the beginning-of-list context, and one node representing the end-of-list distractor task. For Experiment 3 these vectors have 18 total nodes (16 list items + 1 beginning-of-list node + 1 end-of-list distractor).

For Experiment 4, the vectors include additional nodes to represent the studied locations (supermarket, specialty store, farmer's market) and dishes (appetizer, main dish, side dish) as well as one node for each of the three goals participants were assigned (free recall, location-focus, or dish-focus). The relevant goal node replaces the end-of-list distractor node because in the experiment the delay between study and recall of each list was filled with a reminder of task goals. Therefore, the vectors for Experiment 4 contain 28 total nodes (18 list items + 1 beginning-of-list node + 3 location nodes + 3 dish nodes + 3 goal nodes).

Memories are stored through a set of associative matrices that represent the associations between items and states of mental context. When an item is studied or retrieved, it activates its associated state of context, and those feature-context associations are stored in M^{FC}. M^{CF} contains associations between a given state of context and each feature and is used when context is used as a cue for a feature. Each associative matrix is the weighted sum of two matrices, one representing pre-experimental associations (M^{FC}_{pre}, M^{CF}_{pre}), and the other representing associations that are formed during the experiment as items are studied (M^{FC}_{exp}, M^{CF}_{exp}). These experimental matrices are initialized with zeros, and as each item is studied, the item-context and context-item associations are updated.

In the version of CMR used in the simulations for Experiment 3, pre-experimental associations in M^{FC}_{pre} were initially set as an identity matrix to represent that prior to the experiment, each item is only associated with its own context. Semantic associations were implemented in M^{CF}_{pre} such that

$$M^{CF}_{pre} = \begin{cases} \delta, \text{if } i = j. \\ \alpha + s_{cf} M^{FF}_{ij}, \text{if } i \neq j. \end{cases} \tag{A.1}$$

Although previous versions of CMR set the associations between each context and its associated item (e.g., when $i = j$) to be 1 and the base activation for all other associations to be zero, Morton and Polyn (2016) found that allowing the activation levels to vary significantly improved model fit to semantic organization. Here, δ represents the weight of the association between each context and its associated item. This parameter is similar to $1 - \gamma$ in Polyn et al.'s (2009a) original CMR model. α represents the baseline level of support that each context has for all other items in recall competition.

M^{FF} represents the pre-existing semantic associations between items and is multiplied by the model parameter s_{cf}, a semantic scaling factor. In this version of the model, the semantic similarities are derived from Word Association Space (WAS; Steyvers et al., 2004). The diagonal of M^{FF} is set to zero, so the strength of the association between each context and its associated item is controlled only by the δ parameter.

In PEPPR, the associative matrices function in the same way. However, the pre-experimental associations were set to different values. Because semantic associations were not the focus of the simulations for Experiment 4, both M^{FC}_{pre} and M^{CF}_{pre} were initialized as identity matrices without any semantic associations (each item has an association of 1 with itself and 0 with all other items). The δ, α, and s_{cf} parameters were not used. Then, associations between each goal node and the features relevant to that goal were added to M^{CF}_{pre}. The strength of this association was determined by ϵ, a model parameter. The association between each location node (supermarket, specialty store, and farmer's market) and the location-focus goal node was set to the value of ϵ. Similarly, the association between each dish node (appetizer, main dish, side dish) and the dish-focus goal node was set the the value of ϵ. This allows the activation of a goal at retrieval to bias the retrieval process towards task-relevant features.

Encoding

Before the model studies any items, context is initialized with a beginning-of-list context that is orthogonal to any of the pre-existing associations. Studying an item from serial position i activates its associated feature representation f_i on the feature layer. The newly activated context $\mathbf{c}_i^{\mathbf{IN}}$ is restricted to be of unit length:

$$\mathbf{c}_i^{\mathbf{IN}} = \frac{M^{FC}\mathbf{f}_i}{||M^{FC}\mathbf{f}_i||}. \tag{A.2}$$

At the start of an session, the experimental associations (M^{FC}_{exp} and M^{CF}_{exp}) are initialized to zero. As each new item is presented, new experimental associations are formed, both between the item's feature representation f_i and the current state of context c_{i-1} (stored in M^{FC}_{exp}) and between the current state of context and the item's feature representation (stored in M^{CF}_{exp}). These associations are formed according to a Hebbian outer-product learning rule for each matrix:

$$\Delta M^{FC}_{exp} = \mathbf{c}_{i-1}\mathbf{f}_i^{\top} \tag{A.3}$$
$$\Delta M^{CF}_{exp} = \mathbf{f}_i\mathbf{c}_{i-1}^{\top}\phi_i,$$

where ϕ_i simulates increased attention to beginning-of-list items, producing a primacy effect, by scaling the magnitude of context-to-feature associations across the list:

$$\phi_i = \phi_s e^{-\phi_d (i-1)} + 1. \tag{A.4}$$

ϕ_s and ϕ_d are model parameters discussed in greater detail by Sederberg et al. (2008).

These newly formed experimental associations are then combined with the pre-experimental associative matrices. For M^{FC}, the balance between the pre-experimental and experimental associations is controlled by γ_{FC}:

$$M^{FC} = (1 - \gamma_{FC}) M^{FC}_{pre} + \gamma_{FC} M^{FC}_{exp} \tag{A.5}$$

For M^{CF}, the weight of the pre-experimental associations is controlled by delta (when $i = j$) and alpha (when $i \neq j$). Therefore, the context-feature associations are updated by simply adding the pre-experimental and experimental associative matrices together:

$$M^{CF} = M^{CF}_{pre} + M^{CF}_{exp} \tag{A.6}$$

After these new associations are formed,[1] context changes, or drifts, to incorporate the context activated by the just-studied item, $\mathbf{c}_i^{\mathbf{IN}}$, by adding the newly activated context to the current state of context, $\mathbf{c}_{i-1}$. To maintain the context vector at unit length, when a new state of context is added to the existing state the two vectors, $\mathbf{c}_{i-1}$ and $\mathbf{c}_i^{\mathbf{IN}}$, must be scaled so their sum has a length of one:

$$\mathbf{c}_i = \rho_i \mathbf{c}_{i-1} + \beta \mathbf{c}_i^{\mathbf{IN}}. \tag{A.7}$$

Where β is a model parameter governing how quickly context changes, and ρ_i is chosen such that $||\mathbf{c}_i|| = 1$:

$$\rho_i = \sqrt{1 + \beta^2 [(\mathbf{c}_{i-1} \cdot \mathbf{c}_i^{\mathbf{IN}})^2 - 1]} - \beta(\mathbf{c}_{i-1} \cdot \mathbf{c}_i^{\mathbf{IN}}). \tag{A.8}$$

Because context is always of unit length it can be thought of as point on the surface of a (hyper)sphere, with β determining how far along the surface of the sphere it travels with each newly

[1]Associations forming before context is updated is consistent with most implementations of CMR and related models, but not with the version of CMR used by Morton and Polyn (2016), where context drifted prior to new associations being formed.

presented item and $\mathbf{c}_i^{\mathbf{IN}}$ determining the direction of travel. β_{enc} represents the distance context travels with each item that is encoded.

The instructions that intervened between study and test were simulated by assuming that any event during the retention interval causes a change in context. Context was therefore updated using equation A.7 with a different drift rate parameter, $\beta_{distract}$, which represents the rate of context drift when a distractor (i.e., additional instruction) occurs.

Encoding in PEPPR with Goal Representations

The encoding period proceeded in much the same way for CMR and the modified PEPPR with goals used in the Experiment 4 simulations. However, in Experiment 4, there were two additional kinds of encoding events: encoding of task goals and encoding of the location and dish items.

At the beginning of the encoding period, after the beginning-of list context is activated but before the model studies any items, the model encodes the assigned goal. The relevant goal node (free recall, location-focus, or dish-focus) is activated, depending on which condition is being simulated. M^{FC} is updated as specified in equation A.5. In this model, M^{CF} is modified in the same way as M^{FC}:

$$M^{CF} = (1 - \gamma_{CF})M^{CF}_{pre} + \gamma_{CF} M^{CF}_{exp} \tag{A.9}$$

where the the balance between the pre-experimental and experimental associations is controlled by γ_{CF}. Context drifts to reflect the goal state in the same way as when the model encodes an item (equation A.7) but with a different drift rate parameter, $\beta_{PEPPR_{encoding}}$, that represents the rate of context drift when the goal is learned.

When the model begins encoding the list, it studies each element of the ingredient-location-dish triplet individually in the same way as items are studied in CMR. First, the representation f_i of the ingredient is activated on the feature layer, and the associative matrices are updated. Then the ingredient activates its pre-existing associations, and context drifts following equation A.7 with the drift rate parameter β_{enc}. The location associated with the ingredient is then studied in the same way, followed by the dish.

Recall

Before recall begins, the beginning-of-list context is reinstated. As Morton and Polyn (2016) report, this mechanism improved model fit over no reinstatement of the beginning-of-list context, and conceptually it represents the tendency for some participants to think back to the beginning of the list and use this event as a cue (Laming, 1999). Context is updated to reflect the beginning-of-list-context:

$$\mathbf{c}_{start} = \rho_i \mathbf{c}_N + \beta_{start} \mathbf{c}_0, \tag{A.10}$$

where c_{start} is the state of context at the start of the recall period and c_0 represents the beginning-of-list context. The rate of context drift is set to be β_{start}, a model parameter specific to the reinstatement of the beginning-of-list context.

From here, the recall period proceeds as a series of retrieval attempts closely following the implementation used by Morton and Polyn (2016). If the model successfully retrieves an item, the model continues onto the next recall attempt until the maximum number of recall attempts have been made (set to be the length of the list), and then the recall process ends. If the model fails to retrieve an item, no further retrieval attempts are made. The probability of stopping recall because of a failure to recall starts low for the first recall attempt and increases exponentially with each output position:

$$P(stop, j) = \theta_s e^{j\theta_r}, \tag{A.11}$$

where θ_s is a parameter which determines the scaling of the exponential function, θ_r is a parameter which controls the rate at which the probability of stopping approaches 1, and j is the output position.

For each recall attempt, context is used to cue retrieval of an item using the M^{CF} associations:

$$\mathbf{a} = M^{CF} \mathbf{c}_t, \tag{A.12}$$

The resulting $\mathbf{a}$ gives the degree of support, or activation, for each item in the list. These activations are then used to assign each item a probability of being selected for recall according to:

$$P(i) = (1 - P(stop)) \frac{\mathbf{a}_i^{\tau}}{\sum_k^N \mathbf{a}_k^{\tau}}, \tag{A.13}$$

where τ is a sensitivity parameter that determines how sensitive the model is to different levels of support. With a high value for τ, the model is much more likely to recall items of high activation than those with low activation; with a low value of τ, less activated items have a greater chance of winning. To ensure the model does not assign a recall probability of zero to any item, each element of **a** is set to a minimum value of 10^{-7}.

Once an item i is recalled, its representation f_i is activated on the feature layer, and item i is recorded as a recall. The item then reinstates its associated context, and context drifts as in equation A.7, where the rate of context drift during recall is set to be β_{rec}, a model parameter. This updated context is then used as a cue for the next recall. The cycle of cue-recall-update context-cue continues until the model fails to recall an item (Equation A.11).

Recall in PEPPR with Goal Representations

In the version of PEPPR with goals used for the Experiment 4 simulations, recall of individual items proceeds in the same way as described above. But two additional steps are also involved in the recall process for this version of the model.

In Experiment 4 instead of a distractor the delay between study and recall was filled with a reminder of participants' assigned goals. In the model, at the end of the encoding period the relevant goal node is re-activated and context drifts as in equation A.7, where the rate of context drift was set to be $\beta_{PEPPR_{retrieval}}$, a model parameter, and the context being reinstated is the context associated with the relevant goal. This step replaces the reinstatement of the beginning-of-list context in equation A.10.

Then, recall begins. As described above, the model continues to attempt recalls until the maximum number of recall attempts have been made or the model fails to recall an item. For this version of PEPPR with goals, the maximum number of recall attempts was set to be 3 times the length of the list, since the model could recall not only the target items, but also the location and the dish for each item. The probability of stopping recall because of a failure to recall is defined using equation A.11. At the first retrieval attempt, the model can retrieve the to-be-remembered ingredients, any of the three locations, or any of the three dishes. Once an ingredient, location, or dish i is recalled, its representation f_i is activated on the feature layer. The retrieved item then reinstates its associated context, and context drifts as in equation A.7, where the rate of context drift during recall is set to be β_{rec}, a model parameter. This updated context is then used as a cue for the next recall. However, the model is restricted to only output the to-be-remembered ingredients. If the model retrieves a location or dish, the same process occurs, but the recalled location or dish is not added to the list of recalled items. Once a location or dish has been retrieved, the model is restricted to not retrieve it again until after an ingredient has been retrieved. Once an ingredient has been successfully retrieved, all locations and dishes are again available to retrieve.

This implementation of PEPPR does not include parameters specific to post-production monitoring and detection of list membership included in the original version of PEPPR (Healey & Wahlheim, 2023).

Model Simulations

For modeling Experiment 3, I attempted to minimize the root-mean-squared deviation (RMSD) between the condition's across-subject average in the actual data and the model's simulated data. There were $k = 13$ free parameters in CMR. For the first round of simulations, the model was fit to average recall probability, temporal bias scores, and semantic lag-CRP curves for each condition using a differential evolution algorithm which ran for 5,000 generations. Recall probability, *lag* = +1 for the temporal bias scores, and *semantic lag* = 1 were given additional weight (5) compared to all other points (1) in calculating the RMSD. The model was also fit to the semantic lag-CRP curves alone using a differential evolution algorithm which ran for 5,000 generations, and *semantic lag* = 1 was given additional weight (5) compared to all other points (1) in calculating the RMSD.

At each generation, 3,000 simulated subjects, each with a different set of parameter values, studied and recalled one list. I ran this entire procedure 5 times for each condition in Experiment 3 for the fits to recall probabilities, temporal bias scores, and semantic lag-CRP curves and 10 times for each condition for the fits to the semantic-CRP curves only. The best fitting parameter values and the RMSD values across the 5 model fits to recall probabilities, temporal bias scores, and semantic lag-CRP curves for each condition are listed in Table A.1. The best fitting parameter values and the

RMSD values for the fits to the semantic lag-CRP curves only are listed in Table A.3. The average parameter values and standard deviations across the 10 model fits to recall probabilities, temporal bias scores, and semantic lag-CRP curves are presented in Table A.2. The average parameter values and standard deviations across the 10 model fits to the semantic lag-CRP curves only are presented in Table A.4. To generate simulated data for the figures, I used the best-fitting parameter sets to simulate recalls for 30,000 simulated subjects per condition (each studying one list).

Fits to the data for the semantic lag-CRP curves of the younger adult condition of the PEERS study were completed in the same way as for the Experiment 3 data except that the model fitting procedure was completed only 5 times. The best-fitting parameter values and RMSD are reported in Table A.5, and the average parameter values and RMSD of the 5 fits are reported in Table A.6.

For modeling Experiment 4, I attempted to minimize the RMSD between the condition's across-subject average in the actual data and the model's simulated data. There were $k = 12$ free parameters in PEPPR. The model was fit to average recall probability, temporal bias scores, and chance-adjusted category clustering scores for location and dish for the free recall condition using a differential evolution algorithm which ran for 2,500 generations. Recall probability, $lag = +1$ for the temporal bias scores, and each of the chance-adjusted category clustering scores were given additional weight (5) compared to all other points (1) in calculating the RMSD.

At each generation, 3,000 simulated subjects, each with a different set of parameter values, studied and recalled 1 list. The parameter values and RMSD from the model fit to the free recall condition are reported in Table A.7. To generate simulated data for the figures, I used this parameter set to simulate recalls for 60,000 simulated subjects per condition (each studying one list) for the location-focus and the dish-focus conditions.

Best-fit Parameter Values for CMR Fits to Recall, Temporal Contiguity, and Semantic Contiguity in Experiment 3

Table A.1 Best-fit parameter values for the fits of CMR to the data of each condition in Experiment 3. The model was simultaneously fit to average recall probability, temporal bias scores, and the semantic lag-conditional response probability curve.

Parameter	Temporal/Temporal	Semantic/Temporal	Temporal/Semantic	Semantic/Semantic
ϕ_s	23.790	10.217	26.771	36.409
ϕ_d	11.839	13.828	4.211	33.870
γ_{fc}	0.998	0.461	0.944	0.324
δ	0.193	9.052	5.447	15.633
β_{enc}	0.334	0.750	0.251	0.965
$\beta_{distract}$	0.003	0.357	0.211	0.678
β_{start}	0.527	0.335	0.974	0.974
β_{rec}	0.995	0.980	0.555	0.691
θ_s	0.006	0.026	0.016	0.018
θ_r	0.481	0.165	0.300	0.223
τ	86.895	42.783	59.112	71.430
α	0.390	8.428	1.794	13.784
s_{cf}	1.928	1.098	13.915	1.925
$RMSD$	0.0853	0.0479	0.0937	0.1573

Average Parameter Values for CMR Fits to Recall, Temporal Contiguity, and Semantic Contiguity in Experiment 3

Table A.2 Average (SD) parameter values for the 5 fits of CMR to the data of each condition in Experiment 3. The model was simultaneously fit to average recall probability, temporal bias scores, and the semantic lag-conditional response probability curve.

Parameter	Temporal/Temporal	Semantic/Temporal	Temporal/Semantic	Semantic/Semantic
ϕ_s	1.298 (16.604)	10.217 (12.038)	12.265 (9.818)	13.395 (8.866)
ϕ_d	2.484 (15.879)	13.828 (8.767)	3.459 (0.299)	4.932 (16.151)
γ_{fc}	0.539 (0.175)	0.429 (0.188)	0.935 (0.009)	0.274 (0.199)
δ	0.193 (4.801)	1.555 (3.757)	2.569 (1.619)	3.407 (4.606)
β_{enc}	0.308 (0.130)	0.382 (0.149)	0.251 (0.021)	0.286 (0.274)
$\beta_{distract}$	0.003 (0.354)	0.074 (0.199)	0.211 (0.197)	0.148 (0.265)
β_{start}	0.527 (0.101)	0.335 (0.146)	0.858 (0.044)	0.647 (0.119)
β_{rec}	0.876 0.044()	0.853 (0.053)	0.458 (0.048)	0.646 (0.136)
θ_s	0.006 (0.008)	0.026 (0.008)	0.016 (0.013)	0.002 (0.006)
θ_r	0.142 (0.117)	0.031 (0.051)	0.040 (0.105)	10.223 (0.125)
τ	73.533 (4.584)	42.783 (19.816)	59.112 (7.761)	36.588 (16.321)
α	0.329 (4.625)	1.285 (3.600)	0.404 (1.012)	3.125 (3.996)
s_{cf}	0.023 (0.811)	0.745 (0.748)	9.721 (1.669)	1.925 (3.404)
$RMSD$	0.1045 (0.0173)	0.0539 (0.0051)	0.0969 (0.0028)	0.1622 (0.0048)

Best-fit Parameter Values for CMR Fits to Semantic Contiguity in Experiment 3

Table A.3 Best-fit parameter values for the fits of the CMR to the semantic lag-conditional response probability curves for each condition in Experiment 3.

Parameter	Temporal/Temporal	Semantic/Temporal	Temporal/Semantic	Semantic/Semantic
ϕ_s	35.672	27.422	1.356	32.210
ϕ_d	25.999	42.983	8.796	42.404
γ_{fc}	0.0256	0.899	0.934	0.229
δ	15.178	19.348	30.410	38.741
β_{enc}	0.094	0.9183	0.032	0.871
$\beta_{distract}$	0.466	0.406	0.181	0.352
β_{start}	0.975	0.992	0.490	0.280
β_{rec}	0.549	0.825	0.938	0.881
θ_s	0.184	0.014	0.030	0.027
θ_r	0.423	0.201	0.051	0.051
τ	19.625	27.253	70.939	56.727
α	12.369	10.269	11.103	5.687
s_{cf}	3.600	13.422	8.978	9.552
$RMSD$	0.0012	0.0016	0.0009	0.0009

Average Parameter Values for CMR Fits to Semantic Contiguity in Experiment 3

Table A.4 Average (SD) parameter values for the 10 fits of CMR to the semantic lag-conditional response probability curves for each condition in Experiment 3.

Parameter	Temporal/Temporal	Semantic/Temporal	Temporal/Semantic	Semantic/Semantic
ϕ_s	22.426 (15.394)	27.018 (14.681)	18.525 (12.703)	25.626 (9.775)
ϕ_d	23.854 (18.199)	30.297 (13.098)	24.278 (13.766)	21.987 (15.391)
γ_{fc}	0.424 (0.254)	0.477 (0.285)	0.666 (0.237)	0.423 (0.195)
δ	19.775 (9.148)	23.806 (7.611)	24.072 (5.9622)	29.359 (8.442)
β_{enc}	0.416 (0.279)	0.449 (0.270)	80.518 (0.316)	0.881 (0.089)
$\beta_{distract}$	0.475 (0.229)	0.585 (0.167)	0.330 (0.281)	0.388 (0.178)
β_{start}	0.436 (0.304)	0.620 (0.232)	0.477 (0.228)	0.531 (0.245)
β_{rec}	0.508 (0.225)	0.594 (0.242)	0.549 (0.313)	0.691 (0.290)
θ_s	0.211 (0.040)	0.028 (0.012)	0.024 (0.008)	0.018 (0.009)
θ_r	0.269 (0.182)	0.132 (0.093)	0.088 (0.067)	0.134 (0.079)
τ	39.975 (27.226)	35.478 (16.750)	54.401 (23.117)	56.515 (22.19)
α	8.706 (3.917)	10.145 (3.792)	10.150 (1.658)	7.063 (3.824)
s_{cf}	5.465 (3.301)	6.530 (4.050)	6.837 (3.534)	9.116 (2.894)
$RMSD$	0.0031 (0.0009)	0.0021 (0.0002)	0.0015 (0.0002)	0.0011 (0.0002)

Best-fit Parameter Values for CMR Fits to PEERS Experiment 1

Table A.5 Best-fit parameter values for the fits of CMR to the semantic lag-conditional response probability curve for the younger adult condition in PEERS Experiment 1.

Parameter	
ϕ_s	32.993
ϕ_d	0.499
γ_{fc}	0.662
δ	25.072
β_{enc}	0.638
$\beta_{distract}$	0.654
β_{start}	0.720
β_{rec}	0.231
θ_s	0.014
θ_r	0.078
τ	51.428
α	1.119
s_{cf}	6.387
$RMSD$	0.0010

Average Parameter Values for CMR Fits to PEERS Experiment 1

Table A.6 Average (SD) parameter values for the 5 fits of CMR to the semantic lag-conditional response probability curve for the younger adult condition in PEERS Experiment 1.

Parameter	
ϕ_s	5.348 (14.579)
ϕ_d	0.499 (14.473)
γ_{fc}	0.180 (0.232)
δ	10.261 (6.920)
β_{enc}	0.200 (0.213)
$\beta_{distract}$	0.334 (0.335)
β_{start}	0.105 (0.251)
β_{rec}	0.003 (0.306)
θ_s	0.002 (0.007)
θ_r	0.022 (0.083)
τ	8.559 (18.650)
α	1.119 (3.651)
s_{cf}	3.057 (3.533)
$RMSD$	0.00110 (0.00009)

Parameter Values for PEPPR with Goals for Experiment 4a

Table A.7 Parameter values for the fit of PEPPR with goal nodes to the data of the free recall condition in Experiment 4a. The model was simultaneously fit to average recall probability, temporal bias scores, chance-adjusted location clustering scores, and chance-adjusted dish clustering scores.

Parameter	Free Recall
ϕ_s	19.195
ϕ_d	15.178
γ_{fc}	0.4253
γ_{cf}	0.921
β_{enc}	0.299
$\beta_{PEPPR_{encoding}}$	0.147
$\beta_{PEPPR_{retrieval}}$	0.338
β_{rec}	0.396
θ_s	0.010
θ_r	0.036
τ	75.090
ϵ	10.222
$RMSD$	0.081