PEPPR: A Post-Encoding Pre-Production Reinstatement Model of Dual-List Free Recall

M. Karl Healey

Michigan State University

Christopher N. Wahlheim

University of North Carolina at Greensboro

Author Note

High-performance computing resources were provided by the Institute for Cyber-Enabled Research at Michigan State University. Correspondence concerning this article should be addressed to M. Karl Healey (khealey@msu.edu) at Michigan State University, Department of Psychology, 316 Physics Road, East Lansing, MI.

Abstract

Recent events are easy to recall, but they also interfere with the recall of more distant, non-recent events. In many computational models, non-recent memories are recalled by using the context associated with those events as a cue. Some models, however, do little to explain how people initially activate non-recent contexts in the service of accurate recall. We addressed this limitation by evaluating two candidate mechanisms within the Context-Maintenance and Retrieval model. The first is a Backward-Walk mechanism that iteratively applies a generate/recognize process to covertly retrieve progressively less recent items. The second is a Post-Encoding Pre-Production Reinstatement (PEPPR) mechanism that formally implements a metacognitive control process that reinstates non-recent contexts prior to retrieval. Models including these mechanisms make divergent predictions about the dynamics of response production and monitoring when recalling non-recent items. Before producing non-recent items, Backward-Walk cues covert retrievals of several recent items, whereas PEPPR cues few, if any, covert retrievals of that sort. We tested these predictions using archival data from a dual-list externalized free recall paradigm that required subjects to report all items that came to mind while recalling from the non-recent list. Simulations showed that only the model including PEPPR accurately predicted covert recall patterns. That same model fit the behavioral data well. These findings suggest that self-initiated context reinstatement plays an important role in recall of non-recent memories and provides a formal model that uses a parsimonious non-hierarchical context representation of how such reinstatement might occur.

Keywords: cognitive control; free recall; metacognition; source monitoring; temporal contiguity

PEPPR: A Post-Encoding Pre-Production Reinstatement Model of Dual-List Free Recall

When we try to recall the past, the most recent experiences are often the first to come to mind. This recency effect appears in many situations, including autobiographical memory (Bahrick, Bahrick, & Wittlinger, 1975; Moreton & Ward, 2010; Rubin, 1982), recall of news stories (Uitvlugt & Healey, 2019), and laboratory tasks such as free recall and item recognition (Murdock, 1962; Murdock & Anderson, 1975). Although having easy access to memories of recent events is often beneficial, it can also create retroactive interference that makes it harder to access more distant memories (Jang & Huber, 2008; McGeoch & McDonald, 1931; Underwood, 1945; Unsworth, Spillers, & Brewer, 2012). Here, we develop a model of a context-based mechanism that may allow people to overcome the pull of recent memories to access distant memories. Our model is the first variant of the Context Maintenance and Retrieval model (Lohnas, Polyn, & Kahana, 2015; Polyn, Norman, & Kahana, 2009) to include a pre-retrieval cue specification mechanism of the sort proposed in the metacognition literature (Goldsmith, 2016; Jacoby, Shimizu, Daniels, & Rhodes, 2005). The mechanism is based on the simple notion that in addition to to-be-remembered items, other aspects of the task, such as task instructions, have context representations that can be used as retrieval cues—an idea that has proved useful in modeling a range of phenomena including primacy (Kragel, Morton, & Polyn, 2015; Morton & Polyn, 2016), serial recall (Logan & Cox, 2021, 2023), source memory (Polyn et al., 2009), emotional enhancement of memory (Talmi, Lohnas, & Daw, 2019), and consolidation (Sederberg, Gershman, Polyn, & Norman, 2011). We evaluate the ability of this and another candidate context-based model to both predict and fit the dynamics of response production in a dual-list free recall paradigm that requires recalling from a list studied *before* a more recent list (Wahlheim, Ball, & Richmond, 2017).

Our approach is inspired by studies showing that subjects can access non-recent memories under conditions of retroactive interference in variants of the list-before-last paradigm (Shiffrin, 1970). In these paradigms, subjects study multiple lists, but instead of recalling from the most recent list, they recall from the list before the last. Subjects in these tasks recall many items from the non-recent list and, despite the potential retroactive interference from an intervening list, produce few inter-list intrusions (Jang & Huber, 2008; Sahakyan & Hendricks, 2012; Unsworth et al., 2012; Wahlheim & Garlitch, 2020; Wahlheim & Huff, 2015; Wahlheim, Richmond, Huff, & Dobbins, 2016; Ward & Tan, 2004). Here, we model data from the similar dual-list free recall task (e.g., Unsworth, Brewer, & Spillers, 2013; Wahlheim, Alexander, & Kane, 2019; Wahlheim et al., 2017). In dual-list free recall, illustrated in Figure 1, subjects study two lists separated by a short break and then recall from either the recent list (i.e., List 2) or the earlier list (i.e., List 1).[1] This is similar to the list-before-last paradigm as subjects are sometimes required to retrieve distant memories (from List 1) while avoiding retroactive interference from recent memories (List 2). The key difference between the two paradigms is that whereas in list-before-last subjects study a series of single lists and after each are asked to recall from the list 2-back in the sequence, in the dual-list paradigm, there is a series of distinct trials each with two lists, one labeled List 1 and the other List 2. Our goal here is to evaluate the ability of two models with different context reinstatement mechanisms to explain how people retrieve non-recent memories from List 1 and prevent intrusions from recent memories in List 2.

Accessing Non-recent Events: Theoretical Mechanisms

Because a primary challenge when facing retroactive interference is to overcome the pull of recent memories, we start with a theory of the recency effect. Many theories assume that retrieved context can be a cue for accessing *recent* memories (e.g., Estes, 1955; Farrell, 2012; Howard & Kahana, 2002; McGeoch, 1932; Mensink & Raaijmakers, 1988; Underwood, 1945). These models generally conceive of context as a set of feature elements distinct from items that can each be active to varying degrees at a given moment in time. The ensemble of activation across context elements is assumed to change gradually across

[1] To avoid using different terminology when discussing list-before-last recall and dual-list recall paradigms, we will refer to the targeted non-recent list as “List 1” and the recent list as “List 2”.

time (e.g., by moving gradually through the feature space or by updating between lists), producing different states of context at different points in time. Items are assumed to form associations with the state of context that prevailed during their presentation. Because context change is gradual, the context state at any moment will be more similar to the context state from the prior moment than states from more distant moments. Therefore, using the current context at the end of a list as a retrieval cue naturally produces the recency effect. Framed in this way, the challenge of recalling from earlier lists is to avoid using current context to *recall* recent items and instead use it to *distinguish* recent from non-recent items (e.g., Jang & Huber, 2008; Lohnas et al., 2015). This distinction may be enabled by back- and front-end control processes that subjects can leverage to strategically regulate their memory accuracy (Burgess & Shallice, 1996; Goldsmith, 2016; Halamish, Goldsmith, & Jacoby, 2012; Jacoby, Kelley, & McElree, 1999; Jacoby, Shimizu, Daniels, & Rhodes, 2005; Johnson & Raye, 2000; Morcom, 2016).

Back-end control processes are assumed to guide report decisions after subjects monitor the source of retrieved representations. This prevents memories from non-target sources (i.e., intrusions) from being output (e.g., source-monitoring; Johnson, Hashtroudi, & Lindsay, 1993). Such processes are assumed to enable selective reporting of memories above a confidence criterion, thus increasing the accuracy of output responses only when metacognitive monitoring is effective (Goldsmith, 2016). Evidence that back-end control can improve memory accuracy has been shown in comparisons of performance on free and forced report procedures. Free report procedures encourage subjects to output only responses that are likely to be accurate, thus encouraging the use of back-end control to filter out lower-confidence retrievals. In contrast, forced report procedures encourage subjects to respond to all retrieval cues, thereby preventing response censoring by back-end control processes. Evidence that back-end control improved memory accuracy is shown when the proportion of correct responses is higher following free than forced report instructions. This improvement has been widely observed in tasks such as recall and recognition of word lists (Koriat & Goldsmith, 1994) and general knowledge (Koriat & Goldsmith, 1996), cued recall of paired associates (Kelley & Sahakyan, 2003), and recall from narrated slide shows (Koriat, Goldsmith, Schneider, & Nakash-Dura, 2001). These findings support the view that memory accuracy is governed by the extent to which subjects can evaluate and control reporting of generated response candidates.

By contrast, front-end control processes are assumed to control production quality before retrieval by specifying the source context from which candidate responses should be produced (Burgess & Shallice, 1996; Moscovitch & Melo, 1997). Framing memory queries in this way is assumed to enable cue-dependent retrieval from long-term memory representations from target contexts (Tulving, 1983; Tulving & Thompson, 1973) and prevent memories from non-target contexts from coming to mind (e.g., source-constrained retrieval; Halamish et al., 2012; Jacoby, Shimizu, Daniels, & Rhodes, 2005). Evidence that front-end control can improve response production quality has been observed in recall tasks where multiple candidates are solicited using a "retrieve and report" procedure. In this procedure, subjects are given a cued recall task and are told to report candidate responses as they come to mind while trying to produce the correct response. The efficacy of front-end control in constraining retrieval to the appropriate source is indicated by the extent to which target responses come to mind first. This procedure has shown that experience-driven reduction of proactive interference leads to more target-list responses coming to mind first (Wahlheim & Jacoby, 2011). In addition, after studying lists with separate deep and shallow encoding instructions, subjects produced more target-list responses first when source information was provided during cued recall (Halamish et al., 2012). These findings converge in showing that front-end guidance of retrieval constraints can improve subjects' ability to target specific context features.

Taken together, the findings from these studies suggest that back- and front-end control processes play complementary roles in providing quality control over reporting and response production, respectively. Below, we describe how these putative mechanisms have been incorporated into memory models relevant for understanding how subjects can selectively retrieve in free recall and discuss their potential roles in the dynamics of retrieving non-recent memories.

Back-end control following monitoring. An example of a post-retrieval monitoring process preceding back-end control was proposed in the classic generate-recognize model. In this model, subjects monitor the source context associated with each generated response. Then, subjects use control processes to output retrievals they recognize as being from the target source and withhold retrievals they do not recognize (Anderson & Bower, 1972; Atkinson & Juola, 1974; Bahrick, 1970; Kintsch, 1970). When attempting to retrieve events from non-recent contexts, monitoring processes could be used to compare the retrieved context of generated candidates with context from the target source *after* retrieval (e.g., Winograd, 1968). Lohnas et al. (2015) developed a model including such metacognitive mechanisms. Their Context Maintenance and Retrieval Version 2 (CMR2) model extended earlier retrieved context models (Howard & Kahana, 2002; Polyn et al., 2009; Sederberg, Howard, & Kahana, 2008) of single-list free recall to various across-list phenomena, including retrieval dynamics from the list-before-last paradigm. In CMR2, each study item activates an associated context representation. This newly activated context representation drives context updating by partially replacing the state of context that prevailed before the item appeared, such that the context representations of the most recent items are most strongly activated. Thus, once the whole list has been presented, end-of-list recency items are most strongly represented in context.

When simulating standard free recall from the *recent* list, CMR2 uses a generate-recognize mechanism. First, it produces candidate retrievals by using the context state from the beginning of recall as a retrieval cue. When recall is immediate, the model assumes that this beginning-of-recall context is identical to the context that prevailed at the end of the study period. (To foreshadow, this assumption is a critical difference between the models we consider here.) Second, the model limits recall to the recent list using a “recognition” process that determines the source of the item by comparing the item’s context with the context used to cue retrieval. If the item is from the recent list, the context similarity will be high. If the item is from an earlier list, the similarity will be lower.

Lohnas et al. (2015) suggested that the same context-derived recency signal that provides direct access to recent items in standard free recall could be leveraged to indirectly access non-recent items. When simulating recall from a non-recent list, CMR2’s production phase is the same as in standard free recall: the beginning-of-recall context cues retrieval from the recent list. But the recognition phase is different: rather than rejecting candidate retrievals when context similarity is *low*, it rejects them when context similarity is *high*, on the assumption that an item with context similar to the beginning-of-recall context is from the recent list. Critically, the context associated with the rejected item is reinstated. Although this reinstated context is recent, it will generally be less recent than the recall context and can therefore cue even less recent items. Following this logic, the model could access List 1 in a dual-list task by engaging in a series of covert retrievals (and rejections) of progressively less recent List 2 items. Once it retrieves an item with a context that is below the similarity threshold, the model assumes it has reached List 1, outputs the item, and begins retrieving and outputting additional items associated with similar contexts using the model’s standard recall mechanisms.

If we think of the context associated with the end of List 2 as a physical location, A, and the context associated with the end of List 1 as another location, B, CMR2’s iterative application of the generate-recognize mechanism is like walking backwards such that one’s feet touch many locations between A and B. We therefore label this the *Backward-Walk* mechanism. Lohnas et al. (2015) fit CMR2 to list-before-last data from Jang and Huber (2008) and showed that this mechanism could account for both the overall level of recall from the target list and the number of intrusions from the most recent list. Note that CMR2 deliberately avoided the use of any kind of list-level representation. By contrast, list representations are common in many other models of memory (Cox & Shiffrin, 2017;

Dennis & Humphreys, 2001; Gillund & Shiffrin, 1984; Jang & Huber, 2008; Osth & Dennis, 2015; Shiffrin & Steyvers, 1997). One of the goals of the present work is to test the claim that lists can be discriminated even in the absence of an explicit list-level representation. In particular, Lohnas et al. (2015) did not test whether its list-representation-free back-end mechanism allowed the model to better account for the data than if a front-end reinstatement mechanism was included instead. In the next section, we discuss the potential advantages of a front-end mechanism over a Backward-Walk mechanism.

Front-end control (cue specification). Neuropsychological, verbal, and computational theories (Atkinson & Shiffrin, 1968; Burgess & Shallice, 1996; Jacoby, Shimizu, Daniels, & Rhodes, 2005; Moscovitch & Melo, 1997) have proposed active pre-retrieval processes that allow subjects to effectively specify retrieval cues. For example, theoretical models of memory confabulation in patients with frontal lobe damage propose that such memory errors partly reflect impaired use of strategic processes engaged before retrieval to constrain memory search until a response candidate is produced (Burgess & Shallice, 1996; Moscovitch & Melo, 1997). This notion has been forwarded as an account of how context-guided retrieval attempts during recognition tasks can lead to qualitative differences in the encoding of unstudied memory foils that impact subsequent memory for those foils (e.g., Jacoby, Shimizu, Velanova, and Rhodes 2005; for a review, see Morcom 2016). Further, the idea that control processes guide search for target memories among distractors in memory space is central to computational accounts of human memory (Atkinson & Shiffrin, 1968). These perspectives converge on the common assumption that cue specification reinstates contextual information associated with the target source by recapitulating encoding processes. Such cue-dependent retrieval then increases access to memories from the target context while preventing intrusions from other contexts from coming to mind (Herron & Rugg, 2003; Morcom & Rugg, 2012). When the goal is to recall non-recent events, a perfectly effective cue specification process would reinstate List 1 context *before* response generation, minimizing the activation of List 2 items and preventing their production.

Note that such a context reinstatement mechanism operates after items have been encoded during the study list but before they are produced as candidate retrievals at the start of the recall period. We therefore refer to this class of proactive control mechanisms as Post-Encoding Pre-Production Reinstatement (PEPPR). Apart from CMR2, most formal simulations of list-before-last and related paradigms have implemented a mechanism similar to PEPPR (e.g., Farrell, 2012; Jang & Huber, 2008; Lehman & Malmberg, 2009). These models propose a list-context representation that is independent of items (a feature they share with many models of standard recall; e.g., Anderson & Bower, 1972; Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Kimball, Smith, & Kahana, 2007; Lehman & Malmberg, 2013; Raaijmakers & Shiffrin, 1981).[2] Prior to attempting to retrieve any items, the PEPPR mechanism attempts to reinstate List 1 context. If successfully reinstated, the list context provides a cue to directly access List 1 item representations. If we again think of moving from the end of the List 2 context to the List 1 context as moving between physical locations A and B, PEPPR allows one to jump from point A to point B such that one's feet never touch any intervening locations. Jang and Huber (2008) showed that a version of PEPPR was able to account for both accurate recalls and intrusions in list-before-last recall. However, a major limitation of most models including a PEPPR mechanism is that they assume that list context *can* be reinstated (e.g., Jang & Huber, 2008; Lehman & Malmberg, 2009), without specifying *how* it is reinstated (Lohnas et al., 2015). As we will show, our implementation of PEPPR extends earlier work by providing a formal model of the reinstatement process.

Adjudicating Between Mechanisms

A key challenge in adjudicating between a model that employs a Backward-Walk mechanism versus one that uses a PEPPR mechanism is that both models make similar

[2] Though, the choice of including an unchanging list context is sometimes described as a convenient but unnecessary simplifying assumption (Lehman & Malmberg, 2009).

predictions about summary scores for correct recalls and intrusions. If an idealized subject tried to recall List 1 items using only a Backward-Walk approach, they would successfully recall many List 1 target items and output far fewer List 2 intrusions. An idealized subject using only a PEPPR approach would enjoy comparable retrieval success. Although the two models make nearly identical predictions for summary scores, they make competing predictions about the dynamics of covert retrievals that are withheld from recall.

The subject engaging in Backward-Walk should initiate memory search by producing and withholding List 2 items based on the high similarity between current and retrieved context. That is, a Backward-Walk does not really overcome the pull of recency; rather, it *uses* recency to produce and withhold many List 2 items on the way to List 1 items. Indeed, Lohnas et al. (2015, p. 357) assume that “With a pause between lists, CMR2 covertly retrieves more intervening-list intrusions before discovering the target list.” In contrast, the subject who proactively reinstates context cues via a PEPPR mechanism should produce few, if any, List 2 items. Here, the pull of recency is avoided altogether by jumping directly back to the List 1 context.

These predictions cannot be tested in standard recall paradigms because those tasks require subjects to report only the items from the target list and covertly withhold candidates from non-target lists. This strict report criterion may therefore obscure estimates of intrusions produced with associated List 2 context. Fortunately, the *externalized* free recall procedure (EFR; e.g., Bousfield & Rosner, 1970; Kahana, Dolan, Sauder, & Wingfield, 2005; Roediger & Payne, 1985) allows one to assess both the production and monitoring processes that lead to the response output patterns observed in standard free recall. In EFR, production processes are assessed by instructing subjects to report all items that come to mind as they attempt to recall from only a target list. Monitoring processes are then assessed by instructing subjects to withhold productions that are not from the target list. Although EFR has mainly been used in studies where subjects attempt to recall the most recent list (Kahana et al., 2005; Unsworth & Brewer, 2010; Unsworth, Brewer, & Spillers, 2010), some studies have used EFR to investigate the effects of age and interpolated retrieval on production and monitoring processes involved in recall from non-recent lists (Wahlheim et al., 2019, 2017; Wahlheim & Garlitch, 2020). To directly compare the Backward-Walk and PEPPR mechanisms, we used EFR data in dual-list free recall from Wahlheim et al. (2017). Although EFR provides more information about which items subjects consider and reject than is available in standard recall, it must be acknowledged that it is possible that other items are produced but never enter conscious awareness and thus cannot be reported. In the next section, we summarize the data and our modeling approach.

Overview of Data and Model

The Wahlheim et al. (2017) Data

We tested competing predictions from the two CMR2 models by modeling data from a dual-list free recall experiment that included EFR instructions (Wahlheim et al., 2017). We will briefly review the method here. The experiment included 30 younger adults (and an older adult group that we do not consider here)[3]. On each trial, subjects studied two 10-word lists. Each word appeared for 1 s followed by a 1 s interstimulus interval. Before the first item of each list appeared, the list name (i.e., List 1 or List 2) appeared for 3 s. After studying List 2, subjects were prompted to recall from either one or both lists using one of the following recall prompts: “List 1”, “List 2”, or “Lists 1 and 2”, that appeared for 3 s. Subjects were instructed to type the words from the list(s) indicated by the recall prompts and to report any other words that came to mind while doing so. For each response, subjects pressed a button to indicate whether they thought it was from a target list (i.e., correct) or not from a target list (i.e., incorrect). After each judgment, they rated

[3] We chose to focus only on the younger adult data; developing and testing the model blind to the older adult data. Our hope is that the model we present here can eventually be used to examine cognitive aging by identifying the aspects of the model that would need to change to account for age-related recall differences.

their confidence in that judgment using a 1 (low) - 3 (high) scale (we do not consider the confidence data here). Each subject completed five study-test trials from each of the three target list conditions (i.e., List 1, List 2, and Both Lists 1 and 2) in pseudo-random order.

We first examined recall of responses classified by subjects as being from the correct list(s). Note that these summary scores are comparable to standard free recall scores when subjects are instructed to only output target list responses. Figure 2B shows the proportion of items reported *and* endorsed as from a target list for each study list in each target-list condition. Correct recall was higher from the target list than the non-target list when subjects recalled from only one list. For those trials, subjects correctly recalled about half the items and reported very few inter-list intrusions. These patterns indicate that subjects could effectively discriminate target from non-target items when the target list was not recent (i.e., List 1 condition).

The Context Maintenance and Retrieval Model Version 2

We examined the mechanisms that enable such targeting of non-recent lists by fitting two versions of the CMR2 model to these data: one that uses a Backward-Walk mechanism (as in Lohnas et al., 2015), and one that uses a PEPPR mechanism that we introduce here. We are primarily interested in the List 1 condition because it requires retrieval of non-recent item representations. Nonetheless, because subjects must encode the lists without knowing which will be the target, we also modeled summary scores in the other conditions. Moreover, because subjects complete multiple trials of dual-list recall during an experimental session and must deal with proactive interference from prior trials, we modeled an entire session of trials, allowing memory representations, and thus interference, to accumulate across trials. Modeling all experimental conditions across a full session forced the model to use a common set of study-phase parameters to fit to all conditions and use only test-phase parameters to fit the differences between conditions. Before describing the mechanisms in detail, we provide a non-technical overview of the aspects of the model that the two versions share. A full formal description is available in the Supplemental Materials.

In CMR2, a feature layer and a context layer are connected by associative weight matrices: an item-to-context matrix and a separate context-to-item matrix. Nodes on the feature layer represent individual items, whereas the corresponding nodes on the context layer represent contextual associates activated by each item. To illustrate how these layers interact during study and recall, we will work through an example using Figure 3 as a schematic representation of the model. For this first example, we will illustrate the steps involved in studying and recalling the items *river*, *money*, and *onion* from most recent list (i.e., standard free recall). Later, we will extend the example to recall from a non-recent list.

Simulating encoding during the study phase. Suppose that the word *river* appeared first during study (see first row of Figure 3). The river node would then become active on the feature layer. This activation would project through the feature-to-context associative matrix to activate the context layer. Conceptually, this context layer activity represents various pre-experimental episodic and semantic associates of the word *river*. For example, seeing that word may cause a subject to imagine a river, think of the word's meaning, or remember when they last swam in a river. Presentation of the next word (second row of Figure 3), *money*, then activates its feature layer representation, which completely replaces the feature layer representation of *river*. That is, feature representations are active on the feature layer only while corresponding items appear. Once *money*'s feature representation is activated, it in turn activates its associated context on the context layer. But unlike the feature layer, context layer activation persists across item presentations such that *money* does not replace *river*.

This process repeats when the next item, *onion*, appears (third row of Figure 3). Notice that although all items are active on the context layer, they are not all equally active. Instead, as each new item appears, the activity of all contexts associated with previous items diminishes. This is illustrated by the size of the images on the context layer in Figure 3. A model parameter, $\beta_{encoding}$, governs the rate at which context activation fades across item presentations. When $\beta_{encoding}$ is close to 1, context changes very quickly such that the context of each newly presented item almost completely replaces the contexts of previous items. By contrast, when $\beta_{encoding}$ is close to 0, context remains largely unchanged when new items are presented. That is, the higher the value of $\beta_{encoding}$, the stronger the potential recency effect. Critically, when each new item appears, its feature representation forms new associations with the current state of the context representation. For example, when *onion* appears, a new association can form between the *onion* item representation and the blended *river/money* context representation, thereby encoding the context in which *onion* was studied.

As described in more detail in the Supplemental Materials, when modeling a session of dual-list recall, separate nodes are devoted to representing the items that will be studied on each trial. Specifically, for each trial the feature and context layers include a node for each List 1 item and each List 2 item, plus an extra elements to represent the change in context believed to take place between lists and between trials (Lohnas et al., 2015; Sahakyan & Kelley, 2002). The degree of this between-list and between-trial context updating is governed by two separate parameters, $\beta_{between_{lists}}$ and $\beta_{between_{trials}}$.

Simulating the recall period. During the recall test, the current state of context is used as a retrieval cue by projecting the start-of-test context through the matrix of context-to-item associations. As described earlier, cuing with the current state of context generates a particular level of support for accessing each item representation. These support values are then used as input to a competitive decision process (here, the decision process was a simple Luce-choice rule; see the Supplemental Material for details). Generally, the start-of-test context best supports retrieval of item representations near the end of the list, so those representations typically win the competition and are the first to be considered for output.

We say *considered* for output because in CMR2, sampling an item representation (i.e., an item winning the competition) is the production phase that occurs before a decision is made about outputting a response. Thus, before the sampled item is output, an editing process allows the winning item to reinstate its associated context representation by projecting the item's representation through the matrix of item-to-context associations. For example, thinking of *onion* would result in the reinstatement of the *onion* context and the blended *river/money* context that was active when *onion* was encoded. This reinstated context is then compared to the current mental context to determine if the sampled item representation was from the target list (i.e., the most recent list). If the similarity of the two contexts exceeds a threshold, the sampled item is output; if the similarity is below the threshold, it is withheld. Regardless of the winning item's fate, the reinstated context is integrated with the current mental context, and the resulting updated context cues the next recall. Incorporating the context of the just-sampled item with the context cue biases the next recall attempt to produce an item representation from a nearby serial position.

Related to the present study, Lohnas et al. (2015) used this generate-recognize process to simulate retrieval dynamics in EFR when recalling from non-recent lists. Their version of the model treats an item representation winning the retrieval competition as equivalent to a subject reporting an item during EFR and treats the post-production output decision based on context monitoring as equivalent to a subject pressing a button to output or withhold the retrieved representation. But when Lohnas et al. (2015) modeled list-before-list recall, there were no available data from an experiment with EFR instructions; therefore, they could not test the specific predictions that we do here. The aspects of the model we have described so far provide the foundation that allows CMR2 to simulate standard recall of the most recent list. Next, we will describe how both the Backward-Walk and the PEPPR variants of CMR2 build on this foundation to initiate retrieval from a non-recent list.

Implementing Backward-Walk

Lohnas et al. (2015) implemented the Backward-Walk mechanism to account for recall initiation in the list-before-last paradigm by modifying CMR2's retrieval monitoring mechanism. Specifically, they modified the model to target a non-recent list by reversing the context-comparison editing process. In their model, retrieved items with *high* similarity to the context cue are recognized as intrusions from the most recent list and are thus withheld. However, the context of each withheld item becomes part of a cue that elicits a subsequent candidate production (i.e., the context updates with each retrieval attempt). When this process is applied iteratively, each successive production is likely to come from an earlier serial position than the last. This backward walk eventually produces a candidate associated with a context that is dissimilar enough from the cue context to surpass the output threshold. At this point, the model assumes the target list context has been reached and reverts to the normal editing process in which retrieval products are output only when their contexts are sufficiently *similar* to the cue context.

Note that the outcome of the editing process can create a distinction between the order in which items are sampled (i.e., win a decision competition and be considered for output) and the order in which sampled items are output (i.e., exceeding the report threshold and being endorsed as a correct recall). The model makes specific predictions about both sampling and output orders when targeting List 1. Regarding output order, Lohnas et al. (2015) showed that the model predicts that the probability of outputting an item from the most recent list (i.e., intruding a non-target item) should increase across output positions and found that the data from Jang and Huber (2008) supported this prediction. However, the model makes very different predictions about sampling order. Specifically, because of the Backward-Walk mechanism, the model predicts that item representations from the most recent list will be sampled early and that items from the target list will be sampled much later, once the Backward-Walk reaches the target list context. Next, we will discuss a mechanism that makes a contrary prediction: that representations from the recent list will be sampled rarely at the outset of the recall period.

Implementing Post-Encoding Pre-Production Reinstatement (PEPPR)

At the heart of the PEPPR mechanism is the assumption that *before* any candidate item representations are sampled, subjects can proactively reinstate contextual features that, during study, had become associated with items from the target list. Other models have implemented a mechanism similar to PEPPR by adopting a hierarchical representational structure in which item representations are associated with a superordinate list-context representation and adding a mechanism that probabilistically reactivates the appropriate list context during recall to provide direct access to items (e.g., Jang & Huber, 2008; Lehman & Malmberg, 2009). However, assuming that non-recent *items* are accessed by first reinstating a non-recent *context* raises the issue of how subjects overcome the interference from recent context representations to reinstate non-recent contexts. That is, it shifts the question from interference among items to interference among contexts.

We address this issue by proposing a general mechanism for directly reinstating non-recent contexts that follows naturally from the principles that govern context retrieval dynamics in CMR2. Our mechanism requires a single simple assumption: context change is driven not only by processing items but also by processing task instructions, such as a prompt to prepare to study List 1 or List 2. That is, processing a task instruction activates a distinct context representation that becomes associated with subsequent item representations. This is similar to how Polyn et al. (2009) modeled the effect of varying encoding operations (e.g., animacy versus size judgments) by assuming that tasks and items activate unique context representations and is related to various ways in which other CMR-inspired models have used context cues to selectively assess items based on task instructions or stimuli characteristics (e.g., Kragel et al., 2015; Logan & Cox, 2021, 2023; Morton & Polyn, 2016; Sederberg et al., 2011; Talmi et al., 2019)

When combined with CMR2's context change and reinstatement mechanisms, the notion that tasks have context representations provides a straightforward account of dual-list free recall. Because context is persistent, items appearing after a task goal can form associations with the task-context representation. Thus, if each list follows a different task instruction (e.g., "prepare for List 1" versus "prepare for List 2"), the resulting distinct task-context representations provide natural list-context tags. During recall, these tags will be reinstated if the recall instructions (e.g., "recall List 1") activate task-context representations similar to those that were activated after subjects read the study instructions (e.g., "study List 1").

We implemented these ideas in the model by adding "List 1" and "List 2" nodes to the feature and context layers. These nodes have no special status and behave like item nodes. That is, reading the instruction to study List 1 activates its node on the feature layer which in turn activates its node on the context layer. This is illustrated in the first row of Figure 4. Once active, the list context element behaves as an item context element—its activity gradually fades as context evolves across the list (rows 2–3 of Figure 4). Note that this fading has two important consequences. First, List 1 items will become associated with the List 1 context when new item-to-context associations are formed. Second, the fading of List 1 context across the list will cause later items to be less strongly associated with the list context.

When List 1 ends and the List 2 study instruction appears, context change continues in the same way. As shown in row 5 of Figure 4, the List 2 context element will be activated and incorporated with context but will not completely displace the List 1 items (as described above, the extent to which context changes between lists is governed by a parameter). As List 2 items appear, the List 2 context element will gradually fade (rows 6–7 of Figure 4).

The final row of Figure 4 shows what happens when the recall period begins. The task of recalling a particular list (here "recall List 1") reinstates the context element of the corresponding list, and it is integrated into the current state of context following exactly the same context evolution process that operated during study but with a new parameter, $\beta_{PEPPR_{reinstatement}}$, determining the extent to which a list-context representation replaces the time-of-test context representation. If $\beta_{PEPPR_{reinstatement}}$ is close to 1, then the List 1 representation will dominate, strongly cuing List 1 representations. If $\beta_{PEPPR_{reinstatement}}$ is close to 0, then little of the List 1 context will be incorporated, weakly cuing List 1 items. Thus, $\beta_{PEPPR_{reinstatement}}$ directly controls the efficacy of proactive context reinstatement, which may vary depending on task and participant characteristics (we return to this point in the Discussion).

Like the Backward-Walk model, the PEPPR model makes specific predictions about both sampling and output orders when targeting List 1. For output order, the PEPPR model predicts that subjects should begin recall with an item from List 1, and that this first output item should come from an early serial position. This prediction is consistent with the probability of first recall patterns of past research (Wahlheim et al., 2019, 2017; Wahlheim & Garlitch, 2020). Moreover, the PEPPER model diverges sharply from the Backward-Walk model in predicting that sampling order should closely follow output order, at least for the first several productions. This is because reinstating the List 1 context representation prior to producing any items allows the model to directly access List 1 item representations without needing to first produce and withhold a long sequence of List 2 items.

Simulation Results

Deriving Predictions of Backward-Walk and PEPPR Models

As discussed above, the Backward-Walk model predicts that when attempting to recall from a non-recent list, subjects should begin by producing items from the most recent list. By contrast, the PEPPR model predicts that subjects should begin recall by reinstating the target-list context and then producing items from that list. To formalize these predictions, we fit both models to the summary scores from Figure 2C (see the

Supplementary Materials for details about model fitting). Simulated data from the best-fitting parameterizations are shown in Figure 2A and Figure 2B for the Backward-Walk and PEPPR models, respectively. Both models fit the data very well, capturing the key finding of near-perfect discrimination of target from non-target items.

We determined the models' predictions about covert response production by recording every item produced when simulating the best-fitting parameterization for each model (i.e., those that won the retrieval competition). We also recorded whether sampled items were rejected after the context-comparison editing process. For a given sampling position, we calculated the probabilities that the item was from List 1 (a correct recall from the target list) or List 2 (an intrusion from the non-target list). Because, like subjects, the model does not always recall all the items; later sampling positions tend to have fewer observations. Therefore, we calculated sampling probabilities for a given position conditional on an item actually having been recalled in that position. Moreover, note that both models were capable of sampling only three types of items: a current-trial List 1 item, a current-trial List 2 item, or a prior-trial item. As a consequence of making our probability measure conditional on some items being sampled, for any given sampling position, the conditional probabilities of the three types of items must sum to one. Therefore, in our figures, we show only the items of interest (current List 1 and List 2 items) and omit prior-trial intrusions. We focused only on trials where the models targeted List 1 because our interest is in understanding the mechanisms underlying recall from non-recent memories.

Figure 5A shows that when trying to recall from List 1, the Backward-Walk model invariably begins by first sampling an intrusion from List 2. In fact, the best-fitting parameterization sampled almost every List 2 item before it sampled any List 1 item. Moreover, the Backward-Walk model almost never produces intrusions from lists prior to the current trial (i.e., all sampled items come from List 1 or List 2 — the conditional sampling probabilities for current trial List 1 items and current trial List 2 items sum to one). By contrast, Figure 5B shows that the PEPPR model tends to begin recall by first sampling a correct recall directly from List 1 and almost never samples a List 2 item until all List 1 items have been sampled. PEPPR does, however, sample intrusions from prior trials (i.e., the conditional sampling probabilities for current trial List 1 items and current trial List 2 items *do not* sum to one). Indeed, after output position 13, PEPPR samples no items from either the current List 1 or List 2, instead sampling only prior-trial intrusions.

In addition to testing the models' predictions about which items are sampled, we also examined their predictions about which sampled items are rejected as intrusions. The Backward-Walk model (Figure 5D) predicts that the first items it samples (which, as we have seen, tend to come from List 2) will almost always be rejected. Once the Backward-Walk model successfully contacts List 1 (i.e., at sampling position 11, which corresponds to having sampled almost all of the items from List 2), the rejection probability drops precipitously. The PEPPR model (Figure 5E) predicts that sampled items, the first of which tend to be from List 1, will almost never be rejected.

Testing the Model Predictions

The simulations above clearly showed that the Backward-Walk and PEPPR models make categorically distinct predictions. The EFR data from Wahlheim et al. (2017) allowed us to evaluate which model's predictions are closer to observed patterns of sampling and rejection. We analyzed the subjects' EFR responses in the same way that we analyzed the models' data. We treated all items that subjects reported as sampled and items that they marked as incorrect as rejected. That is, whereas in the model we can directly measure which items are sampled, for subjects we infer sampling by treating all responses they produced during the EFR procedure as "sampled", allowing us to estimate their sampling probabilities. The resulting sampling and rejection probabilities appear in Figure 5C and F. Panel C shows that subjects almost always began recall by sampling from List 1, which was more consistent with PEPPR than Backward-Walk model predictions. Also supporting PEPPR's predictions, subjects' conditional sampling probabilities for the current List 1 and the current List 2 do not sum to one, indicating that, like PEPPR, they occasionally sampled intrusions from sources other than the current trial. Turning to the rejection probability data (Panel F), subjects rarely rejected the first-sampled items, but the sampling probability increased steadily across positions. These data are not consistent with predictions of either Backward-Walk, which predicts a sharp *decrease* in rejection probability at sampling position 12, or PEPPR, which predicts a near-zero rejection probability for all positions.

Overall, these results provide stronger support for the PEPPR model than the Backward-Walk model. Caution is warranted, however, for two reasons. First, although the PEPPR model predictions match the behavior of subjects more closely than the prediction of the Backward-Walk model, the PEPPR predictions are far from perfect. For example, whereas subjects initiated recall by sampling a List 1 item with about 80% probability and a List 2 item with about 10% probability, the PEPPR model initiates by sampling from List 1 with about 55% probability and from List 2 with zero probability. Further, whereas subjects show a gradual increase in the probability of rejecting a sampled item, the PEPPR model predicts an almost 0% rejection probability for all outputs. Second, the models were never fit to sampling probabilities, rejection probabilities, or correct recalls as a function of serial position. Instead, the predictions were derived from each model's best-fitting parameterization to the summary scores in Figure 2. That is, we have shown that the parameterization of the Backward-Walk model that provides the best-fit to summary scores does not naturally predict subjects' sampling and rejection behavior. Although it seems intuitively unlikely, we must rule out the possibility that the Backward-Walk model could successfully simulate the sampling and rejection behavior if fit directly to the data. Therefore, we attempted to directly fit each model to subjects' estimated sampling probabilities (i.e., the data from Figure 5C), rejection probabilities (i.e., the data from Figure 5F), and summary scores (i.e., the data from Figure 2C). To ensure we explored the parameter space of each model sufficiently to find a good-fitting parameter set, if one exists, we iteratively applied a differential evolution genetic algorithm (see the Supplemental Materials for details).

The results of these simulations are shown in Figure 6. The right column reproduces the subjects' data for direct comparison. The left column shows that the Backward-Walk model could not fit the data: Simulated Backward-Walk subjects showed no bias toward sampling from List 1 (i.e., the target list) over List 2 at any output position and sampled far too few current-trial items overall (Panel A). The middle column shows that the PEPPR model fared much better: Simulated PEPPR subjects precisely paralleled actual subjects' high level of sampling from List 1 and List 2 (Panel B) and also very closely matched subjects' gradual increase in rejection probability across output positions (Panel E). If we focus only on items that the model endorsed as correct recalls from the target list, PEPPR again outperforms Backward-Walk. As a reminder, when targeting List 1, subjects successfully recalled 47.7% of List 1 items and intruded 2.7% of List 2 items (i.e., the first two bars of Figure 2C). The best-fitting version of Backward-Walk successfully recalled 10.0% of List 1 and intruded 10.3% of List 2 (i.e., it was unable to distinguish the two lists). By contrast, PEPPR successfully recalled 40.4% of List 1 items and intruded 13.2% of List 2 items. Although PEPPR provided a much closer fit of the subjects' data than Backward-Walk did, when compared to subjects, PEPPR recalled slightly fewer List 1 items and made considerably more List 2 intrusions. This is not an inherent limitation of PEPPR, as Figure 2B shows the model can precisely fit these summary recall measures when considered in isolation. Nonetheless, we were not able to achieve as close a fit to these data points when simultaneously fitting conditional sampling and rejection probabilities.Future work should explore whether this excess of List 2 intrusions is a genuine replicable limitation of the PEPPR model.. If replicable, multiple factors could contribute to this lack of perfect fit. One possibility is that subjects employ different, non-PEPPR strategies on some trials. For example, although the current simulations suggest subjects do not consistently use a Backward-Walk mechanism, it is possible they use that mechanism on a small sub-set of trials.

As we have seen, PEPPR predicts that when trying to recall a non-recent list, the first item that comes to mind should be an item from that list and not an intrusion from the most recent list. But PEPPR also makes clear predictions about which *item* within the target list should be recalled first. As a final test of the PEPPR model, we will evaluate these predictions by examining Probability of First Recall curves which show the probability of initiating recall at each serial position. Figure 7B shows PEPPR's *predicted* PFR curve in the List 1 condition. It is important to note that these are not model fits, instead they are predictions derived from the best-fitting parameterizations of the models fit to data from Figures 2 and 6—the model was never fit to the PFR. As can be seen, PEPPR predicts that subjects should start recalling from the beginning of the list. This is because the list representation is most strongly associated with the first item (see the discussion of Figure 4). By contrast, the Backward-Walk model (Figure 7A) predicts that recall will be initiated from the last item of the target list, because that is the first item reached when walking backward in time. As seen in Figure 7C, the data agree closely with the PEPPR predictions.

In summary, although predictions of the Backward-Walk model are fully consistent with observed summary scores, the model fails to predict covert sampling patterns even when fit directly to the data. By contrast, the PEPPR model predicted summary scores comparably to the Backward-Walk model while also correctly predicting qualitative patterns of covert sampling and performing quite well when fit directly to those retrievals.

Discussion

Understanding how people overcome the pull of recent memories to recall non-recent events has been a theoretical challenge. We addressed this challenge here by developing a formal model of dual-list free recall including context-based mechanisms inspired by the memory modeling and metacognition literatures. We compared two retrieved-context models that used different mechanisms to produce and output items from a non-recent list. The first model, which we called Backward-Walk, used a back-end mechanism to control the output of items generated from progressively less recent positions (Lohnas et al., 2015). The second model, which we called PEPPR, used a front-end control mechanism to specify retrieval cues in the service of producing non-recent list items and preventing recent-list intrusions from coming to mind. We tested divergent model predictions against data from an experiment in which subjects externalized covert response sampling. The Backward-Walk model's predictions diverged sharply from subject's actual behavior especially for covert recalls. By contrast, the PEPPR model's predictions were in closer agreement with the data even for the critical covert recalls. When we moved beyond predictions and attempted to fit the models directly to the relevant data, PEPPR preformed quite well, but the Backward-Walk model failed. This model comparison highlighted the necessity of considering response sampling dynamics to precisely characterize the context-based mechanisms underlying recall of non-recent events.

Relation to Other Models of Recalling Non-recent Events

Many models of recall of non-recent events assume that item representations are subordinate to list-context representations that can be reinstated via pre-retrieval cue specification (e.g., Jang & Huber, 2008; Lehman & Malmberg, 2009). By contrast, most retrieved-context models have eschewed those notions for parsimony (Lohnas et al., 2015) and instead assume that end-of-study context serves as an initial cue to bootstrap access to non-recent item representations (i.e., the Backward-Walk; but see Polyn et al., 2009). The current simulations show that neither approach is completely correct. Consistent with the view that accessing non-recent items does not require list-context representations, Lohnas et al. (2015) demonstrated that a Backward-Walk model can account for summary scores in standard free recall. However, without a list-context representation, the Backward-Walk model incorrectly predicts that subjects should generate and reject many recent-list items before accessing non-recent-list items. This failed prediction suggests that the ability to reinstate a list-context representation may be necessary to directly access earlier lists and avoid sampling recent-list intrusions (cf. Jang & Huber, 2008).

The present simulations made the key contribution of showing that including a list-context representation does not have to add complexity to the model's representational structures. Previous models implemented list- and item-context representations using hierarchical structures. Those models also included processes for reinstating each kind of context during recall. By contrast, the model including the PEPPR mechanism affords no special status to list-context representations. Both list and item contexts are elements on the same vector representation that are activated by external stimuli following the same rules. Consequently, list and item contexts change across the study period in the same way. In this sense, our model is a hybrid of Lohnas et al. (2015) and Jang and Huber (2008) that embeds list- and item-context reinstatement processes and preserves the parsimony of the Lohnas et al. (2015) model. That is, our simulations show that the basic mechanisms of context dynamics implemented in retrieved context models (Howard & Kahana, 2002; Polyn et al., 2009; Sederberg et al., 2008) are sufficient to account for the recall of both recent and non-recent memories without postulating new types of representations or retrieval mechanisms (for a similar point regarding serial recall, see Logan & Cox, 2023).

Relation to Other Paradigms and Future Directions

Although we focused here on dual-list free recall, the computational principles embodied in the PEPPR model may be relevant to other interference-based paradigms and to research on group and individual differences in episodic memory associated with context processing.

List-before-last recall. Dual-list free recall is similar to the list-before-last paradigm (Shiffrin, 1970). However, one key difference is relevant to how a PEPPR mechanism would operate. In dual-list free recall, each trial includes two labeled study lists and test cues with the same labels. Providing labels supports cue specification, as the model assumes that study labels activate unique list-context representations and test prompts using those labels support the reinstatement of those representations. By contrast, the list-before-last paradigm does not include list labels and requires participants to track list order during study. Nonetheless, a PEPPR mechanism could still operate if different context representations are active during each list. Such representations could emerge from self-initiated strategies, like subvocal rehearsal of list numbers. Representations could also emerge incidentally. For example, off-task thoughts, such as thinking about future plans while waiting for a study list could become part of a distinguishable start-of-list context. This would allow subjects to reinstate that context at the start of recall. The mnemonic consequences of such study behaviors could be assessed via thought reports and overt rehearsals (Rundus, 1971; Ward & Tan, 2004) and externalized covert recalls with thought reports before the first retrievals at test.

A second key difference is that whereas the dual-list task presents both lists before recall, the list-before-last task requires recall after each list. Some have argued that interpolated recall induces context change that isolates lists, thereby counteracting retroactive interference (e.g., Jang & Huber, 2008; Pastötter, Schicker, Niedernhuber, & Bäuml, 2011; Sahakyan & Kelley, 2002). Lohnas et al. (2015)'s CMR2 model incorporated this notion by allowing larger context changes between lists separated by interpolated retrieval and was able to simulate the interpolated retrieval effects. One advantage of our approach in taking CMR2 as our starting point is that our model inherits CMR2's ability to account for such effects.

Another variable that may influence the efficiency, and possibly the applicability, of PEPPR is retention interval. If the delay between List 1 and List 2 is held constant, while the delay between List 2 and recall is lengthened, it likely becomes more difficult to distinguish the two lists (e.g., Winograd, 1968). It is not immediately clear whether PEPPR would be able to accommodate such a result as one might expect the cue to retrieve a particular list would reinstate list context regardless of the retention interval. One possibility is that after a delay the relative similarity between the two lists may increase (G. D. A. Brown, Neath, & Chater, 2007; Nairne, Neath, Serra, & Byun, 1997; Sederberg, Miller, Howard, & Kahana, 2010, smiliar to ratio rule in delayed versus continual distractor recall). This would place higher demands on a front-end control mechanism to specify the target search set, which may be problematic for populations whose cognitive control mechanisms operate with lower efficacy.

Serial recall. Serial recall of items in their original order requires initiating retrieval with a non-recent item—the least recent item in the most recent list. The list-context representation in the PEPPR model could be extended with minimal modifications to account for some aspects of serial recall. Indeed, Logan and Cox (Logan, 2021; Logan & Cox, 2021, 2023) have simulated many aspects of serial recall using their CMR-inspired Context Retrieval and Updating model in which a list context element was activated at the outset of study and allowed to fade as context evolves across subsequent items, such that incorporating list context into the recall cue would produce a primacy gradient in the serial position curve (for a commentary see Osth & Hurlstone, 2023). This proposal is reminiscent of extant serial recall models with hierarchical representations (e.g. Farrell, 2006, 2012). It is also consistent with evidence that free and serial recall rely on similar mechanisms (e.g., Bhatarah, Ward, & Tan, 2006; F. M. Brown, Neft, & LaJambe, 2008; Ward, Tan, & Grenfell-Essam, 2010). Moreover, modeling of reaction time distributions has provided evidence for start-of-list reinstatement (Laming, 1999; Osth & Farrell, 2019; Osth, Reed, & Farrell, 2021).

Standard free recall. A similar application of the PEPPR mechanism may also help retrieved context models account for the primacy effect in standard free recall. Indeed, in an attempt to provide a better fit to the primacy effect, some versions of CMR have incorporated a mechanism that reactivates a beginning-of-list context during recall (Kragel et al., 2015; Morton & Polyn, 2016). There is evidence that the primacy effect arises from recall being initiated from primacy positions and is modulated by subject-controlled factors. For example, subjects more often initiate recall from the first item when they attempt to recall all the items (Tan, Ward, Paulauskaite, & Markou, 2016; Ward & Tan, 2019). This suggests that subjects could use PEPPR to strategically initiate recall to satisfy task demands. Such findings may be accounted for by assuming that list-context representations are more strongly associated with earlier than later list items due to context updating across study events. List context representations could also potentially be used to simplify or supplement CMR2's generate/recognize mechanism for rejecting intrusions by using a mismatch in list context representations to reject prior-list intrusions.[4]

Other tasks that require access to specific memories. Thus far, we considered situations where recent items interfere with recall of non-recent items. But recency is not the only impediment to accessing more distant memories. We propose that a PEPPR mechanism could potentially support access to task-relevant memories in various situations.

For example, classic context dependent memory effects, in which memory is better if study and test occur in the same physical context (Godden and Baddeley 1975, but see Murre 2021), may partly depend on the extent to which a PEPPR mechanism can reinstate the study context. Consistent with this possibility, Smith (1979) showed that context-dependent retrieval requires mentally reinstating study context at test. They eliminated the benefits of physically reinstating the study context at test by asking subjects to mentally reinstate the original study context before being tested in a different physical context. Similarly, encouraging subjects to think back to extra-experimental experiences between two study lists can lead to greater forgetting of the first list, presumably because by inducing mental context change that isolates lists (Delaney, Sahakyan, Kelley, & Zimmerman, 2010; Pastötter & Bäuml, 2007; Sahakyan & Kelley, 2002). Models including a PEPPR mechanism could formally characterize the roles of context-based processes in these and similar effects of subject-controlled context reinstatement.

[4] We thank an anonymous reviewer for making this suggestion.

PEPPR may also be relevant when items must be encoded and accessed according to semantic context. For example, when exemplars from various categories are studied then retrieved, the same neural activity that differentiated categories during study can be reinstated to guide memory search. The degree of such reinstatement predicts the extent of category clustering during recall (Morton et al., 2013; Polyn, Natu, Cohen, & Norman, 2005), implicating a potential neural signature of PEPPR applied to semantic context. Another kind of semantic context is the value associated with items. Studies of value-based encoding show that items assigned higher values during study are better recalled (Castel, 2008) and recognized with more contextual detail (Hennessee, Castel, & Knowlton, 2017). This could reflect self-directed context reinstatement to a "valuable" item category governed by a PEPPR mechanism. Indeed, a similar suggestion was made to account for variations in context-based retrieval across study items assigned different values (Stefanidi, Ellis, & Brewer, 2018). Note also that when reinstating semantic context, subjects may repeatedly engage PEPPR after the initial recall attempt when the retrieved context does not automatically cue target memories. This could occur, for example, when transitioning recall to a new category after exhausting recall from another category or when subsequently recalling high-value items from more distant temporal contexts (i.e., non-adjacent input positions).

Individual and group differences. Individual differences in context processing have long been linked to differences in memory ability (Healey, Crutchley, & Kahana, 2014; Kane, Conway, Hambrick, & Engle, 2007; Sahakyan, Abushanab, Smith, & Gray, 2014; Unsworth & Spillers, 2010b), particularly in situations that require overcoming interference (Burgoyne & Engle, 2020; Hasher & Zacks, 1988; Healey, 2016; Kane & Engle, 2000; Unsworth & Engle, 2007). There have been few attempts to formally model these differences (Healey & Kahana, 2014; Lehman & Malmberg, 2013). Variation in the efficiency of the PEPPR mechanism may contribute to such associations. For example, people with better cognitive control abilities organize their retrieval from target contexts more efficiently. This has been shown in associations between working memory capacity and both context-based recall (Healey et al., 2014; Spillers & Unsworth, 2011; Wahlheim et al., 2019) and the selection of encoding and retrieval strategies (Unsworth, Miller, & Robison, 2019; Unsworth & Spillers, 2010a).

The principles of PEPPR may also provide insight into group differences. Differences in context processing have often been inferred from studies examining populations with impaired memory abilities, such as schizophrenia patients (Polyn et al., 2015), adults with high subclinical schizotypy symptoms (Sahakyan & Kwapil, 2016, 2018), children and adolescents at risk for schizophrenia (İmamoğlu et al., 2022), adolescents with ADHD (Gibson, Healey, & Gondoli, 2019), and older adults (Wahlheim & Huff, 2015; Wingfield & Kahana, 2002). The memory impairments experienced by these group could partly reflect a dysfunctional PEPPR mechanism. For example, consider age-related episodic memory deficits. Such deficits may reflect impaired self-initiated processing (Craik, 1983; Craik, Luo, & Sakuta, 2010), including strategic encoding (Dunlosky & Hertzog, 1998) and retrieval (Unsworth, 2017). Furthermore, these deficits may reflect a common underlying dysfunction in the processing of certain kinds of context processing (Farrell, 2012; Healey & Kahana, 2016; Wahlheim & Huff, 2015). PEPPR may provide a common thread to connect these perspectives by suggesting that pre-retrieval processing plays a central role in maintaining memory acuity under conditions of interference (also see, Morcom, 2016). A fruitful approach could be to correlate subject-level estimates reflecting the efficacy of PEPPR, such as $\beta_{PEPPR_{reinstatement}}$, with latent variable estimates of component memory and control processes associated with context processing and control over its reinstatement.

Conclusion

In summary, our current modeling efforts suggested that a front-end cue specification mechanism that uses task goals to drive context updating can promote direct retrieval of non-recent item representations under conditions of retroactive interference. We extended a retrieved-context model by adding a PEPPR mechanism that directly reinstates list-context representations. Importantly, including this mechanism allowed the model to fit precise measures of sampling and output dynamics that could not be accounted for when retrieval initiation was assumed to follow an iterative generate-recognize process starting from the end of the most recent list. Our model holds promise as it may account for cue dependent retrieval phenomena involving front- and back-end control processes that vary across populations with cognitive control differences. The present discovery represents the first step towards a larger enterprise aimed at integrating theoretical concepts, methods, and analytical approaches from the often disconnected but clearly complementary literatures on computational memory modeling and metacognition.

Acknowledgments

We thank Lynn J. Lohnas for many helpful discussions and Mitchell G. Uitvlugt for help producing the figures.

Declarations

Funding

This material is based upon work supported by the National Science Foundation under Grant No. 1848972.

Conflicts of Interest/Competing Interests

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethics Approval

Approval was obtained from the ethics committee of Michigan State University.

Open Practices

Data and code used in this manuscript are available at
https://osf.io/kd9c5/?view_only=711e57606acc46b0a78bcfd5866776e6

References

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psycholgical Review*, *79*(2), 97–123. doi: 10.1037/h0033773

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology.* San Francisco: Freeman.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–105). New York: Academic Press.

Bahrick, H. P. (1970). Two-phase model for prompted recall. *Psychological Review*, *77*(3), 215–222. doi: 10.1037/h0029099

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, *104*(1), 54–75. doi: 10.1037/0096-3445.104.1.54

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. doi: 10.1037/0278-7393.32.2.215

Bousfield, W. A., & Rosner, S. R. (1970). Free versus uninhibited recall. *Psychonomic Science*, *20*, 75–76. doi: 10.3758/BF03335608

Brown, F. M., Neft, E. E., & LaJambe, C. M. (2008). Collegiate rowing crew performance varies by morningness-eveningness. *Journal of Strength and Conditioning Research*, *22*(6), 1894–1900. doi: 10.1519/jsc.0b013e318187534c

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–576.

Burgess, P., & Shallice, T. (1996). Confabulation and the Control of Recollection. *Memory*, *4*(4), 359–412. doi: 10.1080/096582196388906

Burgoyne, A. P., & Engle, R. W. (2020). Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science*, *29*(6), 624–630. doi: 10.1177/0963721420969371

Castel, A. D. (2008). The adaptive and strategic use of memory by older adults: Evaluative processing andvalue-directed remembering. In A. S. Benjamin & B. H. Ross (Eds.), *The psychology of learning and motivation: Vol. 48. Skill and strategy in memory use* (pp. 225–270). San Diego: Elsevier Academic Press.

Cox, G. E., & Shiffrin, R. M. (2017). A dynamic approach to recognition memory. *Psychological Review*, *124*(6), 795–860. doi: 10.1037/rev0000076

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *302*(1110), 341–359. doi: 10.1098/rstb.1983.0059

Craik, F. I. M., Luo, L., & Sakuta, Y. (2010). Effects of aging and divided attention on memory for items and their contexts. *Psychology and Aging*, *25*(4), 968–979. doi: 10.1037/a0020276

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. doi: 10.1037/0033-295X.112.1.3

Delaney, P. F., Sahakyan, L., Kelley, C. M., & Zimmerman, C. A. (2010). Remembering to forget: The amnesic effect of daydreaming. *Psychological Science*, *21*(7), 1036–1042. doi: 10.1177/0956797610374739

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review*, *108*, 452–478. doi: 10.1037/0033-295x.108.2.452

Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: What is the role of strategy production? *Psychology and Aging*, *13*, 597–607.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression.

Psychological Review, *62*(3), 145–154. doi: 10.1037/h0048509

Farrell, S. (2006, Nov). Mixed-list phonological similarity effects in delayed serial recall. *Journal of Memory and Language*, *55*(4), 587–600.

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

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. doi: 10.1037/abn0000393

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1–67. doi: 10.1037/0033-295X.91.1.1

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. doi: 10.1111/j.2044-8295.1975.tb01468.x

Goldsmith, M. (2016). Metacognitive quality-control processes in memory retrieval and reporting. In J. Dunlosky & S. K. Tauber (Eds.), *The oxford handbook of metamemory* (pp. 357–385). New York, NY: Oxford University Press.

Halamish, V., Goldsmith, M., & Jacoby, L. L. (2012). Source-constrained recall: Front-end and back-end control of retrieval quality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(1), 1–15. doi: 10.1037/a0025053

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In *Psychology of learning and motivation* (pp. 193–225). Elsevier. doi: 10.1016/s0079-7421(08)60041-9

Healey, M. K. (2016). The encyclopedia of adulthood and aging. In S. K. Whitbourne (Ed.), *The encyclopedia of adulthood and aging* (chap. Interference). New York: Wiley-Blackwell. doi: 10.1002/9781118521373.wbeaa067

Healey, M. K., Crutchley, P., & Kahana, M. J. (2014). Individual differences in memory search and their relation to intelligence. *Journal of Experimental Psychology:*

General, *143*(4), 1553–1569. doi: 10.1037/a0036306

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. doi: 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. doi: 10.1037/rev0000015

Hennessee, J. P., Castel, A. D., & Knowlton, B. J. (2017). Recognizing what matters: Value improves recognition by selectively enhancing recollection. *Journal of Memory and Language*, *94*, 195–205. doi: 10.1016/j.jml.2016.12.004

Herron, J. E., & Rugg, M. D. (2003). Retrieval orientation and the control of recollection. *Journal of Cognitive Neuroscience*, *15*(6), 843–854. doi: 10.1162/089892903322370762

Howard, M. W., & Kahana, M. J. (2002). When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, *46*(3), 85–98. doi: 10.1006/jmla.2001.2798

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. doi: 10.1016/j.scog.2022.100241

Jacoby, L. L., Kelley, C. M., & McElree, B. D. (1999). The role of cognitive control: Early selection versus late correction. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories on social psychology* (pp. 383–400). New York, NY: Guilford Press.

Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. G. (2005). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review*, *12*(5), 852–857. doi: 10.3758/bf03196776

Jacoby, L. L., Shimizu, Y., Velanova, K., & Rhodes, M. G. (2005). Age differences in depth of retrieval: Memory for foils. *Journal of Memory and Language*, *52*(4), 493–504.

doi: 10.1016/j.jml.2005.01.007

Jang, Y., & Huber, D. E. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 112–127. doi: 10.1037/0278-7393.34.1.112

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. doi: 10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (2000). Memory, brain and belief. In D. L. Schacter & E. Scarry (Eds.), (pp. 35–86). Cambridge, MA: Harvard University Press.

Kahana, M. J., Dolan, E. D., Sauder, C. L., & Wingfield, A. (2005). Intrusions in episodic recall: Age differences in editing of overt responses. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(2), 92–97. doi: 10.1093/geronb/60.2.P92

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. Towse (Eds.), (pp. 21–48). New York: Oxford University Press.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 336–358. doi: 10.1037/0278-7393.26.2.336

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language*, *48*(4), 704–721. doi: 10.1016/S0749-596X(02)00504-1

Kimball, D. R., Smith, T. A., & Kahana, M. J. (2007). The fSAM model of false recall. *Psychological Review*, *114*(4), 954–993. doi: 10.1037/0033-295X.114.4.954

Kintsch, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory* (pp. 331–373). Academic Press New York.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, *123*(3), 297–315. doi: 10.1037/0096-3445.123.3.297

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*(3), 490–517. doi: 10.1037/0033-295x.103.3.490

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology*, *79*(4), 405–437. doi: 10.1006/jecp.2000.2612

Kragel, J. E., Morton, N. W., & Polyn, S. M. (2015). Neural activity in the medial temporal lobe reveals the fidelity of mental time travel. *Journal of Neuroscience*, *35*(7), 2914–2926. doi: 10.1523/jneurosci.3378-14.2015

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

Lehman, M., & Malmberg, K. J. (2009). A global theory of remembering and forgetting from multilple lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 970–988. doi: 10.1037/a0015728

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

Logan, G. D. (2021). Serial order in perception, memory, and action. *Psychological Review*, *128*(1), 1–44. doi: 10.1037/rev0000253

Logan, G. D., & Cox, G. E. (2021). Serial memory: Putting chains and position codes in context. *Psychological Review*, *128*(6), 1197–1205. doi: 10.1037/rev0000327

Logan, G. D., & Cox, G. E. (2023). Serial order depends on item-dependent and item-independent contexts. *Psychological Review*. doi: 10.1037/rev0000422

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. doi: 10.1037/a0039036

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, *39*(4), 352–370. doi: 10.1037/h0069819

McGeoch, J. A., & McDonald, W. T. (1931). Meaningful relation and retroactive inhibition. *The American Journal of Psychology*, *43*(4), 579–588. doi: 10.2307/1415159

Mensink, G.-J. M., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review*, *95*(4), 434–455. doi: 10.1037/0033-295X.95.4.434

Morcom, A. M. (2016). Mind over memory: Cuing the aging brain. *Current Directions in Psychological Science*, *25*(3), 143–150. doi: 10.1177/0963721416645536

Morcom, A. M., & Rugg, M. D. (2012). Retrieval orientation and the control of recollection: An fmri study. *Journal of Cognitive Neuroscience*, *24*(12), 2372–2384. doi: 10.1162/jocn_a_00299

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

Morton, N. W., Kahana, M. J., Rosenberg, E. A., Sperling, M. R., Sharan, A. D., & Polyn, S. M. (2013). Category-specific neural oscillations predict recall organization during memory search. *Cerebral Cortex*, *23*(10), 2407–2402. doi: 10.1093/cercor/bhs229

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. doi: 10.1016/j.jml.2015.10.002

Moscovitch, M., & Melo, B. (1997). Strategic retrieval and the frontal lobes: Evidence from confabulation and amnesia. *Neuropsychologia*, *35*(7), 1017–1034. doi: 10.1016/s0028-3932(97)00028-6

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

Murdock, B. B., & Anderson, R. E. (1975). Encoding, storage and retrieval of item information. In R. L. Solso (Ed.), *Information processing and cognition: The loyola symposium* (pp. 145–194). Hillsdale, New Jersey: Lawrence Erlbaum and Associates.

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). doi: 10.1098/rsos.200724

Nairne, J. S., Neath, I., Serra, M., & Byun, E. (1997). Positional distinctiveness and the ratio rule in free recall. *Journal of Memory and Language*, *37*(2), 155–166. doi: 10.1006/jmla.1997.2513

Osth, A. F., & Dennis, S. (2015). Sources of interference in item and associative recognition memory. *Psychological Review*, *122*(2), 260–311. doi: 10.1037/a0038692

Osth, A. F., & Farrell, S. (2019). Using response time distributions and race models to characterize primacy and recency effects in free recall initiation. *Psychological Review*, *126*(4), 578–609. doi: 10.1037/rev0000149

Osth, A. F., & Hurlstone, M. J. (2023). Do item-dependent context representations underlie serial order in cognition? commentary on logan (2021). *Psychological Review*, *130*(2), 513–545. doi: 10.1037/rev0000352

Osth, A. F., Reed, A., & Farrell, S. (2021). How do recall requirements affect decision-making in free recall initiation? a linear ballistic accumulator approach. *Memory & Cognition*, *49*(5), 968–983. doi: 10.3758/s13421-020-01117-2

Pastötter, B., & Bäuml, K.-H. (2007). The crucial role of postcue encoding in directed forgetting and context-dependent forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(5), 977. doi: 10.1037/0278-7393.33.5.977

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 287–297. doi: 10.1037/a0021801

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. doi: 10.1080/13546805.2015.1031372

Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science*, *310*(5756), 1963–1966. doi: 10.1126/science.1117645

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

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

Roediger, H. L., & Payne, D. G. (1985). Recall criterion does not affect recall level or hypermnesia: A puzzle for generate/recognize theories. *Memory & Cognition*, *13*(1), 1–7. doi: 10.3758/bf03198437

Rubin, D. C. (1982). On the retention function for autobiographical memory. *Journal of Verbal Learning and Verbal Behavior*, *21*(1), 21–38. doi: 10.1016/S0022-5371(82)90423-6

Rundus, D. (1971). An analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*(1), 63–77. doi: 10.1037/h0031185

Sahakyan, L., Abushanab, B., Smith, J. R., & Gray, K. J. (2014). Individual differences in contextual storage: Evidence from the list-strength effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 873–881. doi:

10.1037/a0035222

Sahakyan, L., & Hendricks, H. E. (2012). Context change and retrieval difficulty in the list-before-last paradigm. *Memory & Cognition*, *40*(6), 844–860. doi: 10.3758/s13421-012-0198-0

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1064–1072. doi: 10.1037/0278-7393.28.6.1064

Sahakyan, L., & Kwapil, T. R. (2016). Positive schizotypy and negative schizotypy are associated with differential patterns of episodic memory impairment. *Schizophrenia Research: Cognition*, *5*, 35–40. doi: 10.1016/j.scog.2016.07.001

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. doi: 10.1037/xge0000401

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. doi: 10.3758/s13423-011-0086-9

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. doi: 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. doi: 10.3758/MC.38.6.689

Shiffrin, R. M. (1970). Forgetting: Trace erosion or retrieval failure? *Science*, *168*(3939), 1601–1603. doi: 10.1126/science.168.3939.1601

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145. doi: 10.3758/BF03209391

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(5), 460–471. doi: 10.1037/0278-7393.5.5.460

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. doi: 10.1037/a0024852

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

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. doi: 10.1037/rev0000132

Tan, L., Ward, G., Paulauskaite, L., & Markou, M. (2016). Beginning at the beginning: Recall order and the number of words to be recalled. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(8), 1282–1292. doi: 10.1037/xlm0000234

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford.

Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373. doi: 10.1037/h0020071

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

Underwood, B. J. (1945). The effect of successive interpolations and proactive inhibition. *Psychological Monographs*, *59*(3), i–33. doi: 10.1037/h0093547

Unsworth, N. (2017). Examining the dynamics of strategic search from long-term memory. *Journal of Memory and Language*, *93*, 135–153. doi: 10.1016/j.jml.2016.09.005

Unsworth, N., & Brewer, G. (2010). Variation in working memory capacity and intrusions:

Differences in generation or editing? *European Journal of Cognitive Psychology*, *22*(6), 990–1000. doi: 10.1080/09541440903175086

Unsworth, N., Brewer, G., & Spillers, G. (2010). Understanding the dynamics of correct and error responses in free recall: Evidence from externalized free recall. *Memory & Cognition*, *38*(4), 419. doi: 10.3758/MC.38.4.419

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Focusing the search: Proactive and retroactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1742–1756. doi: 10.1037/a0033743

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. doi: 10.1037/0033-295X.114.1.104

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. doi: 10.1177/1747021819847441

Unsworth, N., & Spillers, G. J. (2010a). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*, *17*(2), 200–205. doi: 10.3758/pbr.17.2.200

Unsworth, N., & Spillers, G. J. (2010b). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, *62*(4), 392–406. doi: 10.1016/j.jml.2010.02.001

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Evidence for noisy contextual search: Examining the dynamics of list-before-last recall. *Memory*, *20*(1), 1–13. doi: 10.1080/09658211.2011.626430

Wahlheim, C. N., Alexander, T. R., & Kane, M. J. (2019). Interpolated retrieval effects on list isolation: Individual differences in working memory capacity. *Memory &*

Cognition, *47*(4), 619–642. doi: 10.3758/s13421-019-00893-w

Wahlheim, C. N., Ball, B. H., & Richmond, L. L. (2017). Adult age differences in production and monitoring in dual-list free recall. *Psychology and Aging*, *32*(4), 338–353. doi: 10.1037/pag0000165

Wahlheim, C. N., & Garlitch, S. M. (2020). Adult age differences in the use of temporal and semantic context in dual-list free recall. *Psychology and Aging*, *35*(1), 143–157. doi: 10.1037/pag0000425

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. doi: 10.1037/pag0000049

Wahlheim, C. N., & Jacoby, L. L. (2011). Experience with proactive interference diminishes its effects: mechanisms of change. *Memory & Cognition*, *39*(2), 185–195. doi: 10.3758/s13421-010-0017-4

Wahlheim, C. N., Richmond, L. L., Huff, M. J., & Dobbins, I. G. (2016). Characterizing adult age differences in the initiation and organization of retrieval: A further investigation of retrieval dynamics in dual-list free recall. *Psychology and Aging*, *31*(7), 786–797. doi: 10.1037/pag0000128

Ward, G., & Tan, L. (2004). The effect of the length of to-be-remembered lists and intervening lists on free recall: A reexamination using overt rehearsal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1196–1210. doi: 10.1037/0278-7393.30.6.1196

Ward, G., & Tan, L. (2019). Control processes in short-term storage: Retrieval strategies in immediate recall depend upon the number of words to be recalled. *Memory & Cognition*, *47*, 658–682. doi: 10.3758/s13421-018-0891-8

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.

doi: 10.1037/a0020122

Wingfield, A., & Kahana, M. J. (2002). The dynamics of memory retrieval in older adulthood. *Canadian Journal of Experimental Psychology*, *56*, 187–199. doi: 10.1037/h0087396

Winograd, E. (1968). List differentiation as a function of frequency and retention interval. *Journal of Experimental Psychology*, *76*(2, Pt.2), 1–18. doi: 10.1037/h0025380

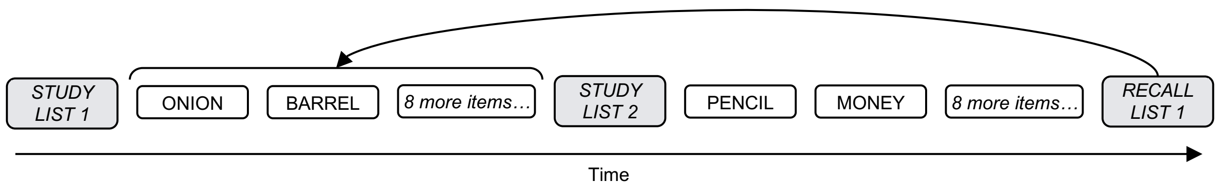

Figure 1. Schematic of the dual-list free recall paradigm from Wahlheim et al. (2017). Subjects studied two 10-item lists that were separated by a short break (i.e., a prompt indicating the start of List 2). At the end of the second list, they were instructed to recall from only List 1, only List 2, or both List 1 and List 2. Here, we focus on the condition where they were instructed to recall non-recent events from List 1, and therefore were tasked to overcome retroactive interference from List 2.

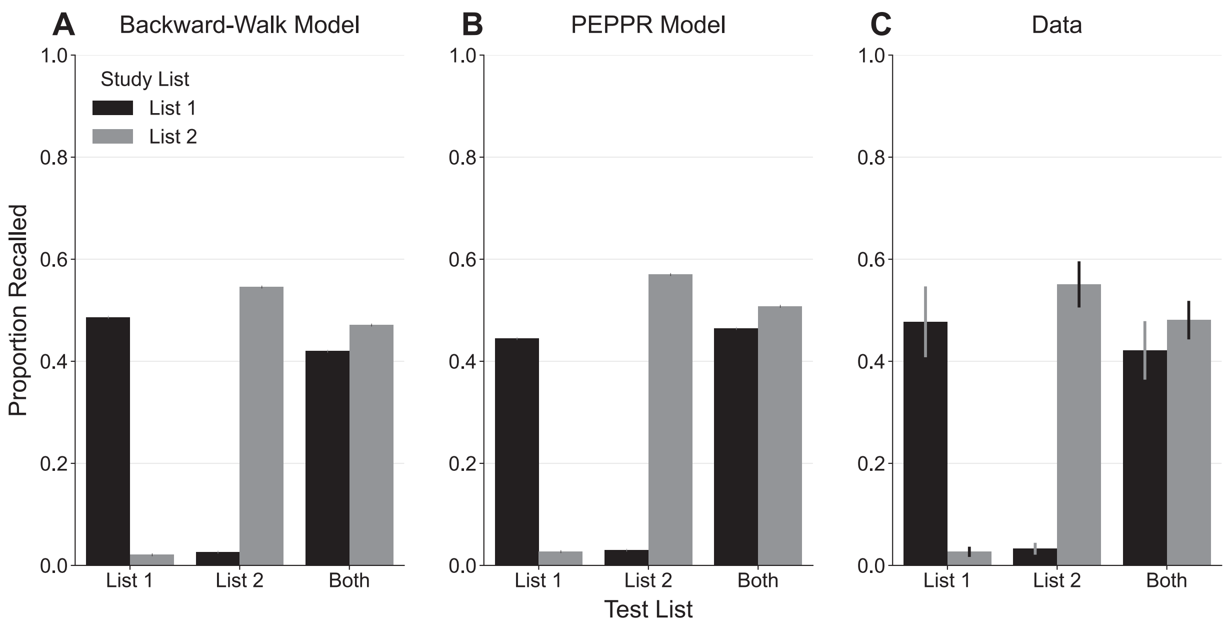

Figure 2. Mean number of items recalled (i.e., were reported *and* endorsed as coming from the target list) as a function of recall test instructions (i.e., recall List 1, List 2, or Lists 1 and 2) and the actual list in which endorsed recalls were originally studied (List 1 or List 2). (A) Best fitting simulated data from the Backward-Walk model. (B) Best fitting simulated data from the Post-Encoding Pre-Production Reinstatement (PEPPR) model. (C) Subject data from Wahlheim et al. (2017). Error bars are bootstrap 95% confidence intervals.

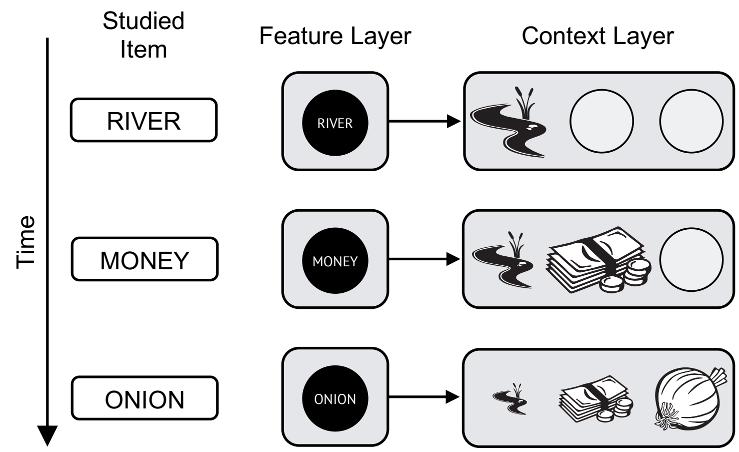

Figure 3. Schematic of the evolution of feature and context representations as items appear, according to CMR2. The feature layer includes a node for each item that will be studied (three in this example), but for simplicity we show only the node for the currently studied item (second column). Corresponding nodes on the context layer (third column) represent the ensemble of contextual associates activated by that item. The two layers are connected by two associative weight matrices (not shown). One matrix stores feature-to-context associations and the other stores context-to-feature associations. When the first item, *river*, appears, its feature layer node is activated. This in turn activates its context representation by projecting through the feature-to-context matrix. This is illustrated by the river icon in the first row of the third column. The context representation of inactive items that have not yet appeared are shown as empty circles in that same row. When the next item, *money*, appears its feature layer node is activated, which then activates and becomes incorporated with its context representation so that both *river* and *money* are active on the context layer. This process then continues with each subsequent study item. Critically, as each new item is activated on the context layer, earlier item context activations decrease (i.e., context changes), illustrated here by the diminishing sizes of pictures in the context layer moving down rows.

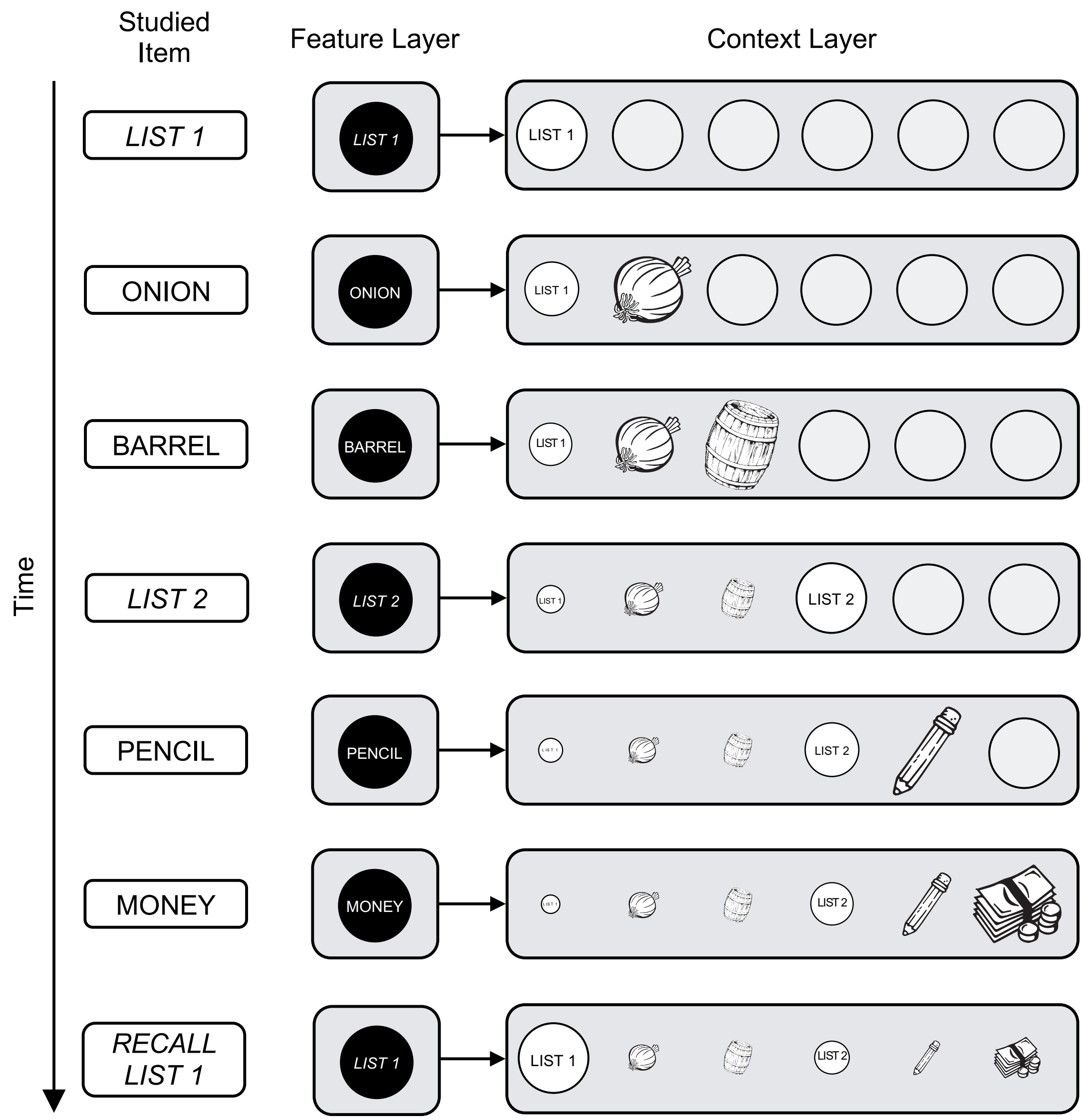

Figure 4. Schematic illustrating how adding task-context representations to CMR2 alters the dynamics of encoding during study and memory search during free recall. All details match the CMR2 illustration in Figure 3, except that layers include representations of study items and task goals (e.g., to study List 1). When the prompt to study List 1 appears, it activates a *List 1* representation on the feature layer, which then activates a corresponding representation on the context layer. These *List 1* context representations and their corresponding item representations become less active with each new study item. When the recall period begins and the prompt to recall List 1 appears, it reactivates the *List 1* feature and context representations, allowing *List 1* context to serve as part of the retrieval cue for the first sampled item. Although all item contexts are active to some degree, the List 1 context is most active.

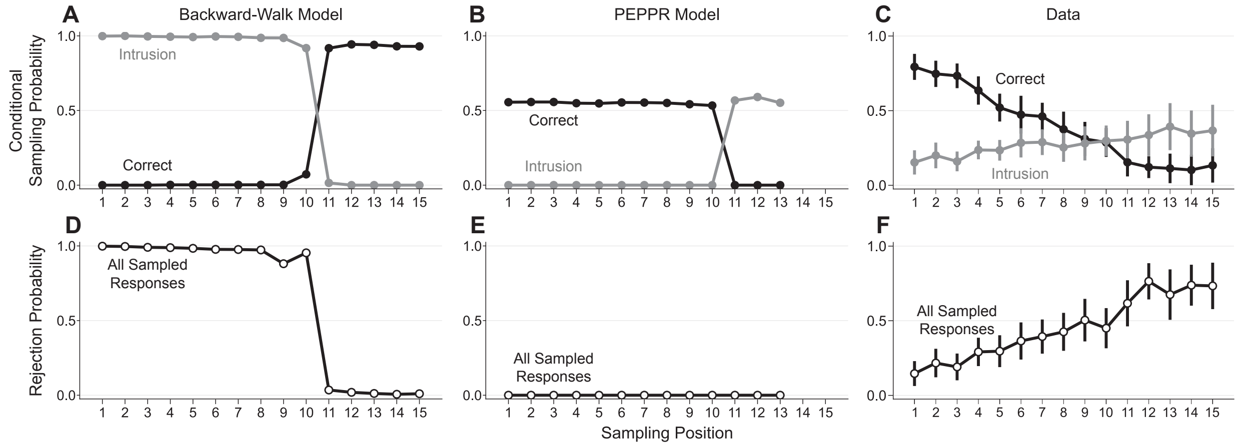

Figure 5. The top row shows predictions from the (A) Backward-Walk and (B) PEPPR models along with (C) subjects' data for the probabilities of sampling correct recalls from List 1 and intrusions from List 2 as a function of sampling position, conditional on an item actually having been sampled in that position. The bottom row shows model predictions from (D) the Backward-Walk and (E) PEPPR models along with (F) subjects' data for the probabilities of rejecting sampled items (i.e., the model/subject believes they are not from the study list) regardless of recall accuracy as a function of sampling position. Error bars are 95% confidence intervals

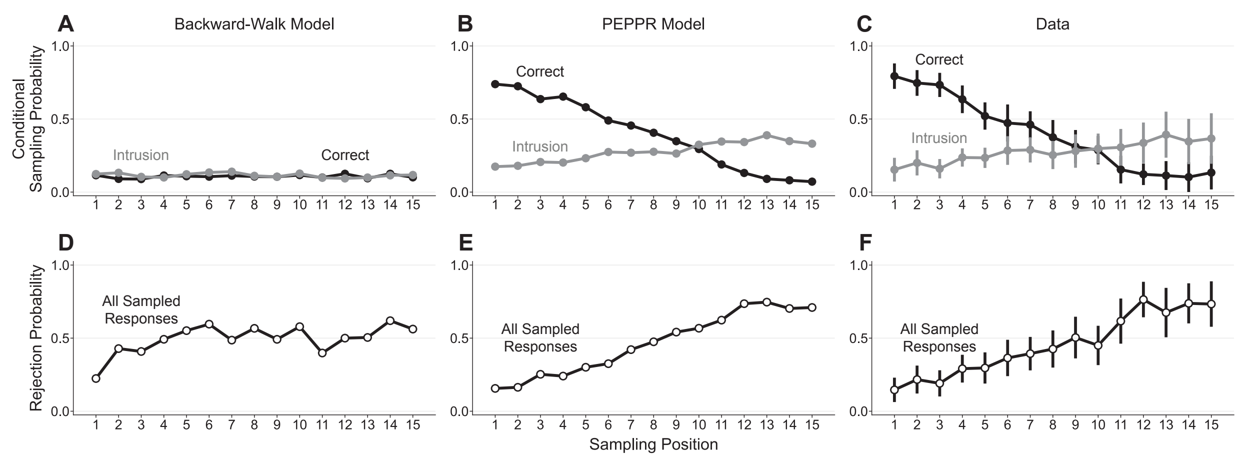

Figure 6. Simulated data from models fit directly to sampling and output data. The top row shows the best-fitting simulated data from the (A) Backward-Walk and (B) PEPPR models along with (C) subjects' data for the probabilities of sampling correct recalls from List 1 and intrusions from List 2 as a function of sampling position, conditional on an item actually having been sampled in that position. The bottom row shows the best-fitting simulated data from (D) the Backward-Walk and (E) PEPPR models along with (F) subjects' data for the probabilities of rejecting sampled items (i.e., the model/subject believes they are not from the study list) regardless of recall accuracy as a function of sampling position. Error bars are 95% confidence intervals.

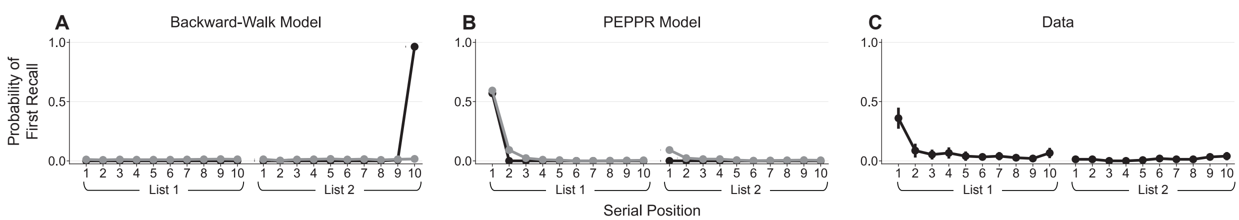

Figure 7. Backward-Walk (A) and PEPPR (B) model predictions for Probability first recall curves, along with actual Probability first recall curves from subject's data (C). In Panels A and B, Black lines are predictions derived from fits to the summary data in Figure 2 and grey lines are predictions derived from to fits from the output profiles in Figure 6. Error bars are bootstrap 95% confidence intervals.